TRAGEDIES of Regional Conflicts & PROMISES of Peacebuilding

Responding to Disruptors and Enablers of MENA Development Pathway

ERF 30th Annual Conference

www.erf.org.eg

On the Political Economy of Conflicts in the Middle East and Africa

Olivier Parent
and Abdallah Zouache

ECONOMIC RESEARCH FORUM 30 years
منتدى البحوث الاقتصادية

2024

On the Political Economy of Conflicts in the Middle East and Africa[1]

Olivier Parent, Department of Economics, Carl H. Lindner College of

Business University of Cincinnati, Cincinnati, OH, 45221

olivier.parent@uc.edu

and

Abdallah Zouache, Sciences Po Lille

9, rue Angelier, 59000 Lille, France

abdallah.zouache@sciencespo-lille.eu

20 December 2023

Abstract

In this research, we aim to identify the main factors that explain the occurrence and intensity of armed conflicts in a specific region, the Middle East and North Africa. We extend the conventional linear Bayesian Model Averaging procedure by incorporating conflict intensity, which is measured across a spectrum of violence levels, departing from the typical binary classification of war or peace. We provide strong evidence that not only demographical, institutional and socio-economic but also, environmental factors must be considered when analyzing conflict intensity. By paying special attention to neighboring states' characteristics, our results reveal that political economy factors, historical legacy, climate and access to natural resources are key in identifying conflict severity. Finally, we show that model averaging predictions for ordered categorical outcomes improve upon the existing out-of-sample conflict prediction techniques.

JEL Codes: O11, O15, C11, C52

Keywords: Conflicts, development, MENA, Africa, Bayesian Model Averaging

1. Introduction

The objective of this article is to provide a new methodology to analyse and predict conflict onset and intensity across countries by focusing on a specific geographical area, the Middle East and

[1] Financial support from the Charles Phelps Taft Research Center and The Excellence Research Committee of the Carl H. Lindner College of Business is gratefully acknowledged. The authors also would like to thank comments from Christopher Adam, David Cobham, Adeel Malik, Mehrdad Vahabi, Saïd Souam and other participants at the Institute for African Worlds seminar, EHESS, Paris, in June 2019, the 2019 workshop of the Research Initiative for the Economics of the Middle East (RIEME), Edinburg, Scotland, the 11th Annual Workshop on "Growth, History and Development", Odense, Denmark, 2020 and the Seminar on Predation state, Conflicts, and Resistance, Sorbonne Paris Nord, in December 2022.

Africa[2], that has been suffering from the world's deadliest conflicts since the fall of the Berlin wall (Gleditsch and Rudolfsen, 2016). Our study faces and responds to different challenges.

Firstly, the nature and propensity of conflicts in North Africa and the Middle East have evolved deeply over the past thirty years, which challenges the ordinal nature of conflicts. Conflicts between states have become very rare whereas internationalized intrastate conflicts (i.e., civil wars with foreign involvement) have consistently increased. As for their types, territorial conflicts persist but purely political conflicts that involve governments and political opposition are also present. The multiplication of armed conflicts in North Africa and the Middle East has been subject to numerous studies (Collier and Hoeffler, 2004). While exploring the main factors influencing the risk of conflict, most comparative studies using cross-country regressions, have often focused on either the onset of civil war (Hegre and Sambanis, 2006) or interstate conflicts (Partell and Palmer, 1999), rarely differentiating between minor and major interstate conflicts. Because armed conflicts vary enormously in size, ranging from disputes with a few fatalities, to massive wars sweeping entire states, the empirical literature has often investigated conflict intensity independently from conflict onset. Focusing on civil conflicts, Lacina (2006) even argues that, despite similarities in the underlying causes, the main determinants of conflict onset reveal little correlation with those for conflict intensity. However, a few studies have guarded against studying those two transitions separately. Sambanis (2004) emphasizes the arbitrary decision surrounding the threshold of fatalities for identifying civil wars. He also underscores the inconsistencies in the number of years of peace that must observed before defining a new conflict. Bluhm et al. (2021) reveal that civil war has never erupted in a civil society that was completely at peace the year before. They show that the cycle of violence often starts with low-intensity conflicts. The ordinal nature of conflicts must therefore be captured over time to fully comprehend the dynamics of armed conflicts.

Secondly, our article will consider a large variety of variables from the literature. Based on a variety of political and economic theories developed to identify the causes of civil war onset (Hirshleifer, 1994), empirical research has focused on a large number of economic, political, social, demographic, and environmental factorsthat can lead to armed conflict (Blattman and Miguel, 2010). One of the most prominent accounts explains conflict onset in terms of greed and grievance (Collier and Hoeffler, 2004). Olson (1965) emphasizes the collective action problem in which individuals fail to cooperate because of conflicting interests. The failure of the social contract between a state and its citizens is also important in the literature. With deteriorating provision of basic services, failure to protect its citizens, and lack of participation in the political decision-making process, the social

[2] We will focus mainly on North Africa, but, as noted below, the conflicts have a tendency to spread so that other African countries are included in our study.

contract breaks down, resulting in a higher risk of violent conflict (Loewe et al. 2021). The literature also points to the causal relationship between climate and conflicts: Burke et al. (2015) review 55 studies on this topic and conclude that deviations from moderate temperature and precipitation patternsincrease the risk of conflict. Climate change issupposed to have a harmful impact on conflicts in Africa because it exacerbatesthe scarcities of natural resources (Mwiturubani and Van Wick 2010). It is also crucial to note that armed conflicts tend to cluster geographically. A number of studies have shown that countries sharing borders with states suffering from instability are more likely to experience conflicts (Ward and Gleditsch, 2002).

Looking across 31 countries in northern Africa and the Middle East over the period 1989-2018, we analyze a comprehensive set of more than 90 potential determinants plus their spatial lags. We introduce spatial lags to control for neighborhood externalities. As model uncertainty is of primary concern when exploring the main factors leading to armed conflicts, we propose a new Ordered Probit Bayesian Model Averaging for longitudinal data that controls for conflict intensity. Even if armed conflicts are often claimed to be too idiosyncratic and complex to allow prediction, the proposed approach surpasses the few existing methodsin terms of out-of-sample prediction accuracy. Our results reveal that colonial legacies in the creation of artificial modern states, lack of economic opportunities, civil liberties and unequal access to renewable resources such as arable land and fresh water are better predictors than measures of religious diversity or economic inequality. Section 2 presents the main challenges when analyzing armed conflicts in the Middle East and North Africa. In Section 3, we provide the theoretical foundationsto justify the potential list of determinants of conflict and we detail the econometric methodology. The main results are analysed in Section 4 along with prediction evaluations. Section 5 concludes and discusses some implications.

2. The nature of conflict in Africa and the Middle East: empirical and conceptual challenges

The dynamics of conflicts tend to be incredibly complex and should not neglect political, economic, and ethnic linkages across state boundaries (Gleditsch, 2007). A few studies have emphasized the link between civil wars and internationalized armed conflicts (Gleditsch, Salehyan, and Schultz, 2008). In the Middle East and North Africa, civil wars combines local and national conflicts in which rebel groups pursue transnational goals (Walter 2017). Internal fighting could spill over into neighboring countries, giving rise to interstate tension, especially in Africa and the Middle East (OECD/SWAC 2022). The ***multidimensional nature of conflicts*** then poses a serious challenge for their measurement and for the identification of the appropriate conflict unit. The ***transnational***

dimension is also essential to understanding why neighboring states are so important in the dynamics of conflict, but it also complicates the definition of our area of study.

In this article, we mostly concentrate on the Arab world, demarcated through the institutional definition of the Arab league, and its contiguous neighbors. Because of lack of information, our sample is made up of 17 out of the 22 members of the Arab league[3] plus Chad, whose official language is Arabic, plus 12 surrounding countries.[4] Nearly half of the members of the Arab League are located in Africa, in the North but also in the Sub-Saharan and eastern parts of the continent. Our period of study starts with the fall of the Berlin Wall in 1989 and ends three decades later in 2018. The longitudinal analysis is paramount for capturing the spillover effects of conflicts taking place in neighboring states. Weak regimes in Africa and the Middle East are more likely to experience instability when sharing borders with states involved in conflicts.

Another complex issue is related to the definition and measurement of armed conflicts. The definition of conflict is usually based on the number of fatalities related to the use of armed force between different organized groups of actors over a year. For each country, the number of fatalities over time is represented in Figure 1 and their location is shown in Figure 2. The logarithmic transformation log(x+1) is used to scale the number of fatalities. Out of the $N=n\times$�=31*30=930 observations, 408 (44%) observations do not contain any fatalities. Around two-thirds of a million fatalities are depicted in Figure 2 over the period 1989-2018. The green-to-red color gradient represents the low-to-high probability of observing a conflict over the period analyzed for each country. Political stability is mainly observed in some of the Arab states of the Persian Gulf such as Qatar, Bahrain, and the United Arab Emirates, but also in North Africa where Morocco and Tunisia have had only a few incidents. Qatar is the only country that did not observe any loss over the entire period. Figure 3 represents the average conflict intensity over time on a scale ranging from 0 to 4 that will be explained in the next section. Most countries suffering from violent conflicts are located in eastern Africa in countries such as Sudan and Somalia but also in the Middle East in Iraq. Conflicts in the Horn of Africa account for more than 50% of all fatalities.

[3]The members of the Arab League are Algeria, Bahraïn, Egypt, Irak, Jordan, Kuwait, Lebanon, Libya, Mauritania, Morocco, Qatar, Saudi Arabia, Somalia, Sudan, Tunisia, United Arab Emirates, and Yemen. Comoros, Djibouti, Oman, Palestine, Syria are left out because of missing data.

[4] Our sample includes Senegal and Mali (sharing a border with Mauritania), Niger, Nigeria, Cameroon, Central African Republic (neighbors of Chad), Eritrea and Ethiopia (bordering Sudan). We did not include countries neighboring South Sudan. Hence, Kenya, Uganda, and the Democratic Republic of Congo are not included even though they were involved in conflicts with Ethiopia, Sudan, and Somalia. In contrast, we have added Burkina Faso, whose recent conflicts have been strongly connected with Mali and Niger. In the Middle East, we have included Israel, Turkey and Iran bordering the Arab countries.

Focusing now on the distinct types of armed conflict, the contested incompatibility that concerns government (50%) is as frequent as that of territorial disputes (50%). Whereas conflicts in Turkey, Israel, and Ethiopia are mostly due to disputed territories, many African countries such as Algeria, Chad, and Somalia but also Iraq and Iran in the Middle East have been suffering from government incompatibilities. Over those thirty years, about twenty percent of armed conflicts involved both incompatibilities. It is important to recognize that international conflicts represent less than 5 percent of all conflicts and involve mostly countries in eastern Africa. More than two-thirds of conflicts are intrastate with a few remaining conflicts being qualified as internationalized intrastate such as the wars in Iraq, Yemen and Somalia.

Complex conceptualization of conflict translates into measures that must encompass various spatial and temporal dimensions. Sambanis (2004) emphasizes that it is nearly impossible to measure civil wars without ad hoc coding rules for war onset and termination. Gersovitz and Kriger (2013) point out the absence of a conceptual definition of civil war since most studies in economics and political science implement a rule-based coding, mainly relying on battle-deaths in the case of civil war.

The question of the number of fatalities remains a core issue. The Uppsala conflict data program (UCDP) requires 25 battle-deaths to define armed conflicts. In contrast, in their historical work on the colonial legacy of conflicts in Africa, Besley and Reynal-Querol (2014) use a threshold of 32 or more battle-related deaths. Gersovitz and Kriger (2013) point out that different criteria can lead to a biased measure of armed conflicts, especially in the case of civil wars, whose occurrence might be over-estimated. In this study, we are mostly interested in analysing the determinants of conflict severity but we recognize the importance of the long-standing literature focusing on the duration and resolution of wars. Exploring the persistence of conflicts would require a careful comparison between the different types of conflicts and the role of combatants, rebellion, and outside parties (Collier, Hoeffler and Soderbom, 2004). Even if duration and frequency of previous conflicts will be key factors in explaining the magnitude of violence, our intent is to identify the best predictors of conflict onset and severity.

In the literature focusing on the determinants of conflicts, the dependent variable usually takes a binary form based on the number of fatalities. When analyzing interstate conflicts, the standard "Correlates of War" project implements a thousand battle-related fatalities for the entire conflict to separate war from non-war (Sarkees 2000). A lower threshold of 25 fatalities is traditionally employed when studying domestic political violence (Melander, 2005). However, from very early on this dichotomous distinction has been criticized for ignoring the difference in magnitude of violence (Duvall, 1976). Melander (2005) differentiates minor conflicts resulting in at least 25 battle-related deaths in one year from civil war causing at least 1000 fatalities. Following Small and Singer (1982),

we assume that conflict-related fatalities effectively capture the severity of conflicts. The Militarized Interstate Dispute (MID) Dataset (Ghosn, Palmer and Bremer, 2004) usesthe number of battle-related deaths to create seven categories of conflict escalation: (0) no fatalities, (1) from 1 to 25, (2) from 26 to 100, (3) from 101 to 250, (4) from 251 to 500, (5) from 501 to 999, and (6) greater than 1000. We combine categories (3), (4) and (5) and propose the following measure based on annual fatalities: (0) peace or absence of conflict, (1) minor violence with less than 25 fatalities, (2) minor armed conflict with fatalities between 25 and 100, (3) intermediate armed conflicts from 100 to 1000 fatalities, and (4) severe armed conflicts with more than 1,000 battle-related deaths per year. This decomposition slightly differs from the UCDP definition of minor and intermediate armed conflicts, which adds a total accumulation constraint of 1,000 deaths during conflicts (Gleditsch et al., 2002).

3. The political economy of conflicts and variable selection

In this section, we first present the main theoretical arguments that are advanced by the literature on the political economy of conflict. We then present the econometric framework used to discuss model uncertainty, model selection, and model averaging.

3.1. Violence, social contracts and the political economy of conflicts in the Middle East and Africa

The key angles adopted in the political economy of social contracts applied to the Middle East and North Africa relate to the natural resource curse (Elbadawi and Selim 2016), the rentier state theory (Beblawi, 1987), neopatrimonialism (Schlumberger 2021), and crony capitalism (Diwan et al. 2019). Access to natural resources (oil, gas, mining) has not only produced specific economic systems but also reoriented political regimes: the standard vision is that African and Middle Eastern states are weak states whose legitimacy and survival is based on authoritarian rule (Schlumberger 2021). Rooted in the clientelist social contract, these policies find their source in colonial history. Colonization has not favoured the implantation of inclusive institutions, especially in Latin America (Acemoglu and Robinson 2012) or Africa (Collier, 2019). Colonization has also played a key role in shaping artificial states (Alesina et al. 2011). Border divisions have led to ethnic diversity, which favors conflicts between groups and makes it more difficult to build a homogeneous state and institutions. In the neopatrimonial political economy system, state failure is the result of strategies of the ruling elites, more obsessed with building socio/economic/political networks of clienteles than building a rational administration (Nugent, 2010: 41). Violence, especially under the form of war, can persist in limited access social order (North et al. 2009), and produces risky and unstable

environments (Bates 2001). Violence prevent the selection of a peace-equilibrium. Conflict is often observed between a ruling elite (usually wealthy) and the organized masses (Acemoglu and Robinson 2000). The threat of a revolution constitutes an incitement for the elite to establish "political settlements" (Acemoglu and Robinson 2000). When conflicts are not avoided, they can take the form of civil war or even repression, forms of political violence that find common roots in poverty and weak institutions (Besley and Persson 2009).

In the MENA region, "welfare states" (Eibl 2020) provide a certain dose of protection to citizens, subsidize the production of certain public goods, and distribute public jobs, the economic bedrock of the middle class, inside an often-corrupted administration (Loewe et al. 2021). In return, individuals enjoy limited participation in political decisions. State-capacity is weak. State capacity should not be limited to military force, as in social orders with limited access, but should enhance its administrative capacity and its ability to negotiate with other actors in society (Fukuyama 2011). Autocratic regimes in the Middle East (Owen 2014) remain legitimate as long as redistribution can take place and violent political conflicts are avoided. As concerns Africa, Collier (2019) states that an active and efficient state is lacking, which impedes the appearance of rules of law, representativeness and political participation (voices and accountability). International aid can effective because democratic assistance programs reduce the risk of conflict (Savun and Tirone 2011) but programs to aid democratization are not always aimed at regime change (Schlumberger 2021).

3.2. The Model

The large variety of theoretical views leads to a vast collection of possible models and makes it very hard to evaluate and promote the most effective policies aimed at reducing political instability. Uncertainty pertaining to the diversity of theories must be addressed by a proper statistical approach. Because economic theory only provides a set of guidelines to identify the proper empirical model, accounting for model uncertainty is fundamental. As described in the previous sections, a large set of structural and proximate factors are certainly relevant when analyzing conflicts. A number of geographical, institutional, cultural, environmental, and socio-economic factors have been suggested by competing theories. Bayesian Model Averaging (BMA) remains the most promising method of accounting for model uncertainty as it directly employs model averaging techniques to identify and estimate parameters of interest. For further reading about BMA, we recommend the seminal papers by Raftery et al. (1997) and Hoeting et al. (1999). Steel (2020) provides comprehensive reviews on model averaging in economics.

Bayesian Model Averaging

A generic representation of an empirical linear regression that analyzes conflict intensity could be the following:

$$y = \alpha \iota_n + X\beta + e, \quad (1)$$

where the *n*-dimensional vector *y* represents the presence of fatalities, *X* designates a set of determinants and *e* is the error term. The scalar α represents the intercept and ι_n is an n-dimensional vector of 1s. The extent to which fatalities *y* can be impacted by the determinants *X* are measured by the marginal effects β. The fundamental question pertaining to the selection of the main determinants *X* remains. Based on competing theories, suppose we are facing *K* possible determinants. Then, we would have the choice between2^K possible combinations of explanatory variables. Even when *K* is moderate, it becomes infeasible to evaluate every model. For instance, we would have to choose between more than 1 million modelsif we had accessto only 20 potential determinants. The following methodology is designed to resolve model uncertainty by constructing estimates that do not rely on a single regression but rather depends on weighted averages across all candidate models. Those candidate models are weighted by their posterior model probabilities based on the following Bayes' theorem:

$$p(M_j|y) = \frac{p(y|M_j)\, p(M_j)}{p(y)}, \quad (2)$$

where M_j represents one of the *j(=1,...,*2^K*)* possible specifications that all seek to explain the dependent variable y. Each model has a prior distribution $p(M_j)$ measuring how likely it is to be correct before observing any data. The function $p(y)$ can be seen as a constant that will cancel out every time two models are being compared. Therefore, the posterior model probability relies mainly on the marginal likelihood $p(y|M_j)$ which can be difficult to evaluate for some specifications. The use of some specific prior structures (see Fernandez et al., 2001) for the linear normal model defined in (1) immediately leads to a closed-form expression for the marginal likelihood but this will not be the case for the non-linear models we will be considering.

The panel data approach used in this study captures not only the variations emerging over time and across space, but also the variation of these two dimensions simultaneously. This is essential in understanding how trends, tendencies and global patterns emerge when analysing conflict intensity. A series of covariates will control for previous lags of conflict intensities, time-invariant effects, and spatial correlation with neighboring countries. In fact, Parent and Zouache (2012) emphasize the importance of geographic neighbors when it comes to analyse the determinants of economic growth across Africa and the Middle East. More recently, Yesilyurt and Elhorst (2017) estimate the strong impact of country spillover effects on the level of military expenditures using spatial dynamic panel data models. In the present study, we rely on a much simpler form of spillover by only introducing an exogenous spatial lag for each determinant. As detailed in the previoussection, we analyze conflict intensity across 31 countries in northern Africa and the Middle East over the period 1989-2018. The different levels of violent conflicts are modelled using the following �(= 5) categories: (0) peace or absence of conflict, (1) minor violence with less than 25 fatalities (2) minor armed conflict with fatalities between 25 and 100, (3) intermediate armed conflicts from 100 to 1000 fatalities, and (4) severe armed conflicts with more than 1,000 battle-related death. The latent variable ��,� represents the propensity of a country *i=1,…,n* at time *t=1,…,T* for entering conflict intensity *j=0,…,J-1*:

� ������,�� l=0,l≠�

��,� = � + ��,�� + ! + ��,�(3)

��,� = j���� < ��,� ≤ ��+1,� = 0,… ,� − 1

where the error term ��,� follows a standard normal distribution with mean zero and the variance ��2 is set to 1 to ensure the model is identifiable. Each coefficient ��� of the *n*×*n* spatial weight matrix *W* is equal to 1 if countries *i* and *l* share a common border and zero otherwise. W is row-normalized so that each spatial lag ���,� represents the average over the neighboring values for the variable ��,� at time *t*. The cut-off points �� are unknown, the *nT×1* vector *Y* of ordered categorical outcomes corresponds to the observed level of conflict intensity and *x* is the *nT*×*k* matrix of covariates. Each response ��,� takes the value *j=(0,…,J-1)*. The probability that country *i* is involved in a conflict of intensity *j* at time *t* corresponds to ��,�,� = �(��,� = �).

The parameters of interest γ and θ are both k-dimensional vectors. Because we are trying to identify the main determinants influencing the propensity of observing conflict-related fatalities, we need to

find an efficient algorithm that would compare all relevant specifications over the entire model space (i.e. all possible combinations). The ��[3] algorithm has been a popular strategy to explore the very large model space (see Masanjala and Papageorgiou, 2008, and Parent and Zouache, 2010, for empirical applications). Because of its specific conjugate priors, Clyde et al. (2011) noticed poor mixing performance when covariates were highly correlated. Two main problems are encountered with probit models that include many explanatory variables. First, to compare models, the marginal likelihood is not available in closed form. Secondly, because the number of potential models is very large, estimating each regression seems prohibitive. Reversible jump Metropolis-Hastings methods (Green, 1995) are typically used to solve simultaneously both issues. Lamnisos et al. (2009) propose a similar transdimensional algorithm that will simultaneously compare specifications across the model space as well as estimate those models if they are deemed relevant. The marginal likelihood is approximated by the Laplace method and the reversible jump samplers are extended to jointly update the model and the latent dependent variables. Because of the ordered categorical outcomes, we rely on data augmentation and simulate the latent variables by integrating out the model parameters β. The algorithm is developed in Appendix A.

4. Results

To make sure the results stay consistent we run four independent Markov Chain Monte Carlo (MCMC) sampling or chains. Each chain was run for 200,000 iterations with a burn-in period of 50,000. With a total sample size of N=31*30=930 observations and 180 explanatory variables including the spatial lags, we have identified around 20,000 unique models for each chain. The top five thousand models account for more than 0.98 of the posterior probability mass. Our results are interpreted via the estimated posterior inclusion probabilities (PIP) and the posterior mean of the model averaging.

4.1. On the determinants of conflicts in Africa and the Middle East

Based on the estimation results presented in Table 2**,** we will first focus on the main determinants that appear consistently in at least 70% of the unique models. Those determinants can be broadly categorized into four groups: institutional and political factors, historical legacies, socioeconomic determinants, and geographical and climate variables.

Political economy variables appear to have strong influence on conflict intensity. Contractual institutions, liberty and respect for property rights are believed to be a necessary step for the

promotion of stability and economic development. As presented in Table 2, academic freedom, polyarchy, corruption and accountability appear with a probability greater than 96% for all four MCMC chains. Government effectiveness and neopatrimonialism have a probability of inclusion greater than 78% for three out of the four chains. Similarly, such a high inclusion probability is observed for rule of law for two MCMC chains. Focusing first on accountability, we find the positive and significant estimate on conflict intensity confirms that accountability can curtail some effective strategies in maintaining order. The aggregate measure of accountability controls for (1) the ability of citizens to exert control over government officials via free and fair elections, (2) the checks and balances exercised by state institutions to oversee the government and separation of powers, and (3) ability for media and civil society to hold government accountable. Although Collier and Rohner (2008) describe a set of mechanisms through which loot-seeking opportunities become less valued as income rises, we still find evidence that democratic states with higher income have suffered tremendously from violent conflicts. Israel and Turkey being prime examples of states with high levels of accountability facing violent conflicts even though instability is more rooted in territorial conflicts than political violence. On the opposite pole, the least accountable countries such as Eritrea, Saudi Arabia, Qatar, Bahrain, and United Arab Emirates have been able to maintain peace for long periods. Democratic indicators that are highly correlated with accountability such as polyarchy and neopatrimonialism are all important predictors for violence. The polyarchy index relies on fundamental democratic principles such as the practice of free and fair elections, the right to run for office, freedom of expression, and the right to form autonomous organizations. Large values of this index are observed in countries like Israel, Turkey and Lebanon along with Sub-Saharan countries such as Senegal, Mali, Burkina Faso and Nigeria. Whereas Burkina Faso has been enjoying until recently relative tranquility, Nigeria and Lebanon have been severely afflicted with violent conflicts. Gulf countries are here again ranked with the lowest value of polyarchy.

One evidence that institutionalist determinants cross each other is that the Fragile State Index (FSI), based on a combination of indicators related to governance, demographic pressures, social cohesion, and economic growth, has an inclusion probability greater than 75% for three out of four MCMC chains.

Our results confirm that in a phase of incomplete democratization, many countries in MENA which had democratic presidential and parliamentary elections lacked solid political institutions (Howard and Roessler 2006), and that led to growing factionalism, triggering ethic violence and armed conflict. The corruption index used here is only associated with embezzlement and lower values indicate a greater level of misappropriation of public funds by government officials. Positive estimatesreinforce the idea that abusing executive power could help authoritarian regimes stay at peace through bribes

and corrupt exchanges. Interestingly, Senegal and Mali put the most value on freedom of academic and cultural expression, which stands out as one of the strongest predictors of stability. Rather than constitutional rights (de jure), this factor controls to what extent actual practices (de facto) of academic and cultural expression are fully respected by authorities. Although the western part of sub Saharan countries enjoys almost no restriction on those civil liberties, censorship and intimidation are more pervasive on the eastern side in war-torn countries such as Eritrea and Sudan.

Politics also counts in terms of military spending and international relations. Firstly, investment in the military sector should not be underestimated in the analysis of conflict dynamics. Indeed, arms imports and the total of armed force personnel have a significant role in shaping political stability with an inclusion probability for two out of the four chains greater than 90% and 75%, respectively. Secondly, the proximity effect seems to work via institutional influence rather than through a purely spatial channel. Indeed, the relationship between international institutions and political stability seems to play an important role as described by the determinant 'affinity with the United States of America' that has an inclusion probability of one for all four MCMC chains. The impact of neighboring countries on own conflict propensity is narrowed to a few institutional factors. Those regional effects are captured by averaging over neighboring observations and seem to be dominated by polyarchy and neopatrimonialism. The spatial lags of those factors also have a strong probability of inclusion but only for half of the MCMC chains. In a third instance, the cultural variables are absent: only the non-Muslim share of population is selected.

The results reveal the importance of a second group of variables related to historical factors. Their influence on the occurrence of conflicts in the Middle East and Africa appears via two channels. A first one could be called a hysteresis factor, in the sense that past conflicts influence the probability of having future conflicts in a country or in a region. This effect is present via the variable "past conflicts in the last three years", which is included in all models, and its spatial lag capturing the effects of past neighborhood conflicts, which has a probability of inclusion greater than 80% for 2 out of the 4 MCMC chains. The second historical set of variables refers to precolonial state development and colonial legacy, as heavily analysed in the literature (Borcan et al, 2018). The results confirm the importance of historical state development in shaping contemporary political stability. The State history score for the 3500 BCE to 1450 CE period has an inclusion probability greater than 70% for half of the MCMC chains. However, recent colonial history should not be underestimated. The Italian and British colonial legacies seem important for half of the chains as well. Artificial borders designed by colonial empires also have a strong influence on political stability as ethno linguistic fractionalization has an inclusion probability of one across all MCMC chains. Regional effects seem also prevalent as the spatial lag of the polarization variable (W*Polarization) is observed

with an inclusion probability of one for all MCMC chains. To a lesser extent, the spatial lag for artificial political borders is also included in more than 70 % of the unique models for two out of the four MCMC chains.

Related to weak institutional factors, a second main obstacle to democratization is thus the incongruity between territory and identity arising from artificial borders designed by colonial powers. Artificial states bore no resemblance to the natural distribution of their indigenous populations (Alesina et al. 2011). Allegiance to a collective agenda is more likely to be weaker in fragmented artificial states. As emphasized in Collier (2001), the impact of ethnic diversity on economic stability depends on the political system and it contributes to economic prosperity in democratic societies. The variable ethno-linguistic fractionalization has a strong and positive effect on the incidence of conflicts. Central African countries such as Chad, Cameroon, Nigeria and Central African Republic are among the most ethnically diverse. They have been suffering from instability unlike the North African countries with the most homogenous ethnic and linguistic groups such as Morocco, Tunisia and Libya. Fractionalization is typically interpreted as the probability that two randomly selected individuals belong to two different groups. In contrast, the ethnic polarization index represents how within-group identity can be ideologically separated from the members of other groups. The maximum value is reached when a state is composed of two groups of equal size. Collier (2001) argues for a non-monotonic relationship between the probability of violent conflicts and ethnic diversity where low risks are only observed for highly homogenous and highly heterogeneous states. Whereas we find no evidence of the direct impact of polarization on state violence, our results suggest positive spillovers from neighboring regions. A state will be less prone to violence if its neighbors have higher levels of polarization. Focusing on the number of battle-deaths, Lacina (2006) emphasizes as well that severity of civil conflicts might be weaker in more polarized states.

Uncertainty regarding the intentions of other actors can increase the risk of political violence. External validations of commitments from democracy-assistance programs can help countries establish democratic governance (Savun and Tirone, 2011). Even if our results do not confirm the importance of international aid, the measure of voting affinity in the United Nations General Assembly (UNGA) with the United States is a strong predictor. Alesina and Dollar (2000) claim that the affinity vote in the United Nations is the main factor explaining the distribution of US aid, even greater than political institutions or economic policy of the recipients. They also reveal that France is giving overwhelmingly to its former colonies and the United States has delivered about one third of its assistance to Egypt and Israel. Our results show that political alignment with the United States, which is often seen as unpopular in North Africa and the Middle East (Carter and Stone, 2015), increases the risk of conflicts.

The third group of factors related to socioeconomic determinants seem to be less important as only a few macroeconomic indicators are correlated with conflict intensity. Unemployment appears to be a major determinant with an inclusion probability of one across all MCMC chains. The employment to population ratio and GDP growth rate are included in more than 75% of the unique models for two out of four MCMC chains. Sociological aspects such as the percentage of female employment remain important with an inclusion probability greater than 90% for two out of the four MCMC chains. Our results confirm that among the dozens of predictors influencing conflict intensity, only a few economic factors related to employment have a strong positive impact on peace and stability. Unemployment has a strong positive impact on conflict intensity, whereas the ratio of employment to population and number of women employed contribute significantly to the promotion of peace.

Lastly and perhaps most importantly, a variety of agricultural and climatic factors appear to be strongly related to water constraints: desalination, water risk in agriculture, volume of surface water entering the country, inflow of water secured through treaty, renewable internal freshwater resources per capita, and percentage of population with access to safe drinking water. Climate and water factors are almost as important as the institutional factors. The extensive research on how climate and conflict are related has led to fierce debates and disagreements (Hsiang and Burke, 2018). The relevant literature converges on the impact of climate change on resource scarcities even if researchers disagree on the mechanism that could translate climate into violence (Gleditsch, 2021). Relying on Collier and Hoeffler's (2004) model of conflict motivated by greed rather than grievance, Gleditsch (2021) argues that greater abundance would not prevent rebellions motivated by a fight over resources. Controlling for historical, institutional and socio-economic factors, our results reveal that access to fresh water, food and fertile land remain major determinants for political stability. The variable measuring the total population with access to safe drinking water has a strong and negative impact on conflict intensity.

While climate change research has raised concern about the loss and damages from extreme weather events, climate-change related variables such as drought risk measures and seasonal variability of available water supply do not seem to have a direct influence on conflict intensity. However, water scarcity can lead to great instability. According to the World Bank (2017), the Middle East and North Africa is the most water-stressed region in the world. Water stress arises when demand for personal, agricultural, and industrial uses outstrips the available level of renewable water resources. Environmental stress overlapping socio-political and economic grievances could increase the risk of tensions. Our results reveal that fierce competition among neighboring states to secure access to water is an important factor of conflict intensity. The importance of transboundary waters, measured by the volume of water entering territories, indicates a high correlation with political instability. Degradation

or depletion could in fact spark conflicts. An integrated approach based on an institutional and legal framework to deal with water resource management is vital to promote peaceful cooperation and development. In fact, our results reveal that water treaties significantly reduce conflict propensity. However, those agreements for managing transboundary water remain rare. Although all countries analysed in this study share at least one aquifer with their neighbors, water management policies are mostly directed toward over-exploiting and depleting aquifers. In fact, another expensive mechanism that helps improve water availability and quality is desalination. However, our results suggest that countries engaged in desalination technologies seem to be more involved in violent conflicts. Even if the volume of desalinated water produced by Saudi Arabia and the United Arab Emirates is greater than all other countries combined, countries like Algeria and Israel are also large producers of desalinated water. Most water policies aim at exploiting the region's fragile aquifers, ignoring the fact that 80 percent of the region's wastewater is lost and could be reused for industrial activities and agriculture (World Bank, 2017).

Although some countries in the Middle East tend to be more industrialized, agriculture remains a key contributor to regional employment (Mwiturubani and Van Wick 2010). Agriculture is known to be the largest consumer of freshwater by far. Our results reveal that countriesrelying on a highly stressed agriculture sector are more peaceful. In fact, in dryland Africa, water scarcity has been a source of social cohesion. With the help of international institutions, regions plagued by severe droughts have been avoiding hostilities. Mauritania, Burkina Faso are facing the highest level of water stress on agriculture and yet have been living relatively at peace compared with some of their neighbors. The regional variable capturing the amount of arable land available in neighboring countries also exerts a strong influence towards peace.

4.2. Conflict prediction and policy implications

We finish our political economy analysis by illustrating how BMA can aid social scientists to make more accurate predictions about future events. Theoretical properties on the predictive performance of BMA can be found in Hoeting et al. (1999). Raftery et al. (1997) show that the quality of forecast always improves when predictions from many models are combined. Although many scholars have acknowledged that predicting international events and trends is a difficult task, various forecasting techniques have been introduced to predict the onset of civil war (Schneider, Gleditsch and Carey, 2011). Most prediction strategies rely on a structural approach such as logistic regression, trying to predict the risk of conflict of a specific geographical unit over time. In addition, some classification techniques based on classification trees and neural network algorithms have been advanced (Beck,

King and Zeng, 2000). We use the top 5,000 models accounting for more than 97% of the posterior model probability to perform predictions. As detailed in Appendix C, we assess predictive accuracy using sampling-based methods for cross-validation prediction. In the binary case, analyzing conflict onsets, Ward and Gleditsch (2002) show a 35 % misclassification. This is slightly better than our proposed model which has a misclassification rate of 37% with five categories (see Appendix C). When analysing the confusion matrix presented in Figure 4, we see that we were able to correctly predict 94 out of the 131 observations related to full scale war (71.8% accuracy). When trying to predict international conflicts, Bleck et al. (1998) successfully predicted 16.7% of conflicts. Ward and Gleditsch (2002) predicted 29 out of 56 international and civil wars (52% accuracy). Gleditsch and Ward (2012), with their best model, were able identify 11 out of 19 conflicts (58%). All those models were working on dichotomous specification of conflict, ignoring conflict severity. The confusion matrix presented in Figure 4 shows that predicting the intermediate level of conflict intensity is a much harder task. We were still able to correctly identify 97 out of the 188 (51.6% accuracy) of the intermediate conflicts. Low intensity remains often confounded with the state of peace which has a correct classification rate of 89%. This unique effort in bringing a large set of factors that control for neighboring effects demonstrates the importance of model averaging to reliably predict conflict intensity.

5. Conclusion

In this study, we propose a Bayesian Model Averaging approach to analyze the incidence and intensity of conflicts in North Africa and the Middle East. We measure conflict intensity by constructing an ordinal outcome based on annual battle-related fatalities. By extending the traditional BMA approach to longitudinal ordered probit models over the period 1989-2018, we exploit the temporal and spatial dimensions to increase model predictability. The proposed procedure allows the selection and estimation of large sets of potential determinantssuch as historical, demographic, socio economic, institutional, and environmental factors while including spatially and temporally lagged covariates.

Although scholars are far from having reached a consensus on how to model conflict onset, core factors are commonly found to be significant. Our results confirm that institutional and economic conditions that favor weak states are the strongest predictors of violence. However, the lack of specificity in many theoretical frameworks often generates tests of partial theories of civil war. This creates problems of omitted variable bias that can affect the validity of estimates. Using a set of 180 potential determinants, our results reveal that the lack of economic opportunities, civil liberties and

unequal accessto renewable resourcessuch asland and fresh water are better indicatorsthan measures of religious diversity or economic inequality.

The precolonial evolution of state institutions seems to have a moderate impact on political stability. However, the important role of colonial legacies that led to ethnic partitioning in the creation of artificial modern states seems to be a strong determinant for conflict intensity. By introducing spatially lagged factors, our results confirm that transnational ethnic linkages represent an important determinant of conflict intensity. Furthermore, many states in Africa and the Middle East did not succeed in developing institutions capable of effectively mobilizing resources and people to guard their territories. Widespread corruption, lack of accountability, and poor governance precipitate violent conflicts. Unlike previous studies, which have analyzed these events mostly in isolation, the proposed moving average procedure used here controls for many factors and reveal some new insights. In fact, our results emphasize that the hardships of climate change, by altering the supply of fresh water and arable land, are likely to add to the burden of food and human insecurity of societies already suffering from weak governments. Falling ground water levels of aquifers shared by many nations and reliance on extensive desalination has accelerated the disparities between demand and water availability. Our results confirm the conventional concern that high pressure arising from a fragmented population could lead to violent conflict over scarce resources. Resilience cannot be achieved if nations develop strategies in isolation. Policy makers should consider climate change an intertwined issue, and recognize that a more efficient access to fresh water across countries will depend on cooperative, sustainable and multidisciplinary international cooperation.

Finally, predictive models seem to perform better with the presence of neighboring factors capturing regional effects. A further investigation needs to be pursued in order to analyze the importance of independencies between countries using spatial econometric specifications such a Spatial Durbin Models within a BMA setting with ordered outcomes. However, our ability to properly assessthe risk of contagion of conflictsrelies on collecting data reflecting intergroup linkages, transnational identity, and shared natural resources.

References

Acemoglu, D. and J.A. Robinson, (2012). *Why Nations Fail. The Origins of Power, Prosperity and Power*, London: Profile Books.

Acemoglu, D. and J. Robinson (2000). "Inequality, Growth and Development. Democratization or Repression ?" *European Economic Review* 44: 683-693.

Alesina, A., Matuszeski, J. and W. Easterly (2011). Artificial states *Journal of the European Economic Association* 9 (2): 246-277.

Aquastat database by Food and Agriculture Organization of the United Nations (FAO) (2019).

Bates, R.J. (2001). *Prosperity and Violence. The Political Economy of Development*. Norton&Co: New York, London.

Beblawi, H. (1987). "The Rentier State in the Arab World", Arab Studies Quarterly, 9 (4), p. 383-398, in Beblawi and Luciani (eds.), *The Rentier State*, London: Croom Helm, p. 49-62.

Besley, Timoty, and Torsten Persson (2009). "Repression or Civil Wars?". *American Economic Review: Papers and Proceedings*. 99 (2): 292-297.

Besley, T., and Reynal-Querol, M. (2014). The legacy of historical conflict: Evidence from Africa.

American Political Science Review, 108(2):319–336.

Blattman, Christopher, and Edward Miguel. 2010. “Civil War.” *Journal of Economic Literature* 48(1): 3–57

Bluhm, Richard, Gassebner, Martin, Langlotz, Sarah, Schaudt, Paul (2021). “Fuelling conflict? (De)escalation and bilateral aid,” *Journal of Applied Econometrics,* 36 (2), pages 244-261

Borcan, O, O. Olsson, and L. Putterman (2018). “State history and economic development: evidence from six millennia”, *Journal of Economic Growth* 23(1): 1-40.

Burke, Marshall, Hsiang, Solomon M, Miguel, Edward (2015). Climate and conflict. *Annual Review of Economics* 7(1): 577–617.

Carter, D. B., & Stone, R. W. (2015). Democracy and Multilateralism: The Case of Vote Buying in the UN General Assembly. *International Organization,* 69(1), pp. 1–33.

Clyde, M., J. Ghosh, and M. Littman (2011). “Bayesian adaptive sampling for variable selection and model averaging”. *Journal of Computational and Graphical Statistics* 20, 80–101.

Collier, P. (2001). “Implications of Ethnic Diversity” *Economic Policy,* 16(32), 128-66.

Collier, P. (2019). “The Political Barriers to Development in Africa”. In Oxford Research Encyclopedia.

Collier, P. and A. Hoeffler, (2004). Greed and grievance in civil war. *Oxford Economic Papers* 56(4): 563–595.

Collier, P., A. Hoeffler and M. Söderbom, (2004). ‘On the Duration of Civil War’, Journal of Peace Research 41(4): 253–273.

Collier, P. and D. Rohner (2008). “Democracy, Development and Conflict”, *Journal of the European Economic Association,* 6, 531-40.

Diwan, I., Malik, A. and I. Atiyas (2019), *Crony Capitalism in the Middle East,* Oxford: Oxford University Press.

Duvall, R. (1976). "An appraisal of the methodological and statistical procedures of the correlates of war project," in Francis W. Hoole and Dina A. Zinnes (eds.) Quantitative International Politics: An Appraisal. New York: Praeger.

Eibl, F. (2020). “Welfare States in the Middle East.” In Cammett, M. and J. Pauline (eds.), *Oxford*

Handbook of Politics in Muslim Societies. Oxford: Oxford University Press.

ElBadawi, Ibrahim, and Hoda Selim. (2016). *Understanding and Avoiding the Oil Curse in Resource rich Arab Economies*. Cambridge: Cambridge University Press.

Fukuyama, F. (2011), *The Origins of Political Order. From Prehuman Times to the French Revolution*. Farrar, Straus and Giroux: New York, Paperback Edition 2012.

Fernandez, C., E. Ley, and M. Steel (2001). "Benchmark priors for Bayesian model averaging". *Journal of Econometrics* 100, 381–427.

Gersovitz, M. and N. Kriger (2013), "What is Civil War? A Critical Review of Its Definition and (Econometric) Consequences", *The World Bank Research Observer*, 28: 159-190.

Gleditsch, Kristian S. (2007). "Transnational Dimensions of Civil War." *Journal of Peace Research* 44(3): 293-309.

Gleditsch, Kristian S., Idean Salehyan, and Kenneth Schultz. (2008). "Fighting at Home, Fighting Abroad: How Civil Wars Lead to International Disputes." *Journal of Conflict Resolution*, 52(4): 479–506.

Gleditsch, Kristian Skrede, and Michael D. Ward. (2000). War and Peace in Space and Time: The Role of Democratization. *International Studies Quarterly* 44 (1): 1–29.

Gleditsch, Kristian Skrede, and Michael D. Ward. (2012). Forecasting is difficult, especially about the future: Using contentious issues to forecast interstate disputes. *Journal of Peace Research,* 50(1):17–31, 2012

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Havard Strand. (2002). Armed conflict 1946-2001: A new dataset. *Journal of Peace Research* 39 (5): 6

Gleditsch, N. P., and Rudolfsen, I. (2016) 'Are Muslim countries more prone to violence?', *Research and Politics*, 3(2), 1-9.

Ghosn, Faten; Glenn Palmer & Stuart A. Bremer, (2004). 'The MID3 Data Set, 1993-2001: Pro cedures, Coding Rules, and Description', *Conflict Management and Peace Science* 21 (2): 133-154

Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82:711 – 732.

Hegre, H. & N. Sambanis, (2006). 'Sensitivity Analysis of the Empirical Literature on Civil War Onset', *Journal of Conflict Resolution* 50(4), 508–535.

Hirshleifer, (1994), “Theorizing about Conflict”. In Hartley, K. and T. Sandler (eds.), *Handbook of Defense Economics*. Vol. 1 : 166-188. Elsevier Science : Amsterdam.

Hoeting,J., D. Madigan, A. Raftery, and C. Volinksy (1999). “Bayesian model averaging: A tutorial”. *Statistical Science* 14, 382–417 (with discussion).

Howard, M.M. and Roessler, P.G. (2006). 'Liberalizing Electoral Outcomes in Competitive Authoritarian Regimes', *American Journal of Political Science*, 50(2): 365-381.

Hsiang, Solomon & Marshall Burke (2018). Conclusion of climate and conflict analysis questioned. *Nature* 554: 587

Lamnisos, D., J. E. Griffin, and M. F. J. Steel (2009). “Transdimensional sampling algorithms for Bayesian variable selection in classification problems with many more variables than observations’’. *Journal of Computational and Graphical Statistics* 18, 592–612.

Loewe, M., Zintl, T. and A. Houdret (2021). “The Social Contract as a Tool of Analysis: Introduction to the Special Issue on “Framing the Evolution of New Social Contract in Middle Eastern and North African Countries”, World Development, In Press.

Masanjala, W. and C. Papageorgiou (2008). “Rough and lonely road to prosperity: a reexamination of the sources of growth in Africa using Bayesian model averaging”. *Journal of Applied Econometrics* 23, 671–82.

Melander, E. (2005) Gender Equality and Intrastate Armed Conflict. *International Studies Quarterly* 49(4):695–714

Mwiturubani, D.A., and Wyk J.A (2010). Climate Change and Natural Resource Conflicts in Africa.Institute for Security Studies Pretoria: South Africa, 277 p.

OECD/SWAC (2022). Borders and Conflict in North and West Africa.OECD: Paris,132 p. https://doi.org/10.1787/6da6d21e-en

North, D., Wallis J. and B. Weingast (2009), *Violence and Social Orders*, Cambridge: Cambridge University Press.

Nugent, P. (2010). “States and Social Contracts in Africa”. *New Left Review*, 63: 35-

68. Olson, M. (1965). *The Logic of collective action*. Harvard: Harvard University

Press.

Owen, R. (2014). *The Rise and Fall of Arab Presidents for Life*. Harvard: Harvard University Press.

282 p.

Parent, O. and A. Zouache (2012). “Geography versus Institutions: New perspectives on the growth of Africa and the Middle East”, *Journal of Institutional and Theoretical Economics*, 168 (3): 488-518.

Partell, Peter J., and Glenn Palmer. (1999). “Audience Costs and Interstate Crises: An Empirical Assessment of Fearon's Model of Dispute Outcomes”. *International Studies Quarterly* 43 (2):389-406

Raftery, A., D. Madigan, and J. Hoeting (1997). “Bayesian model averaging for linear regression models”. *Journal of the American Statistical Association* 92, 179–91

Sarkees, Meredith Reid. 2000. The Correlates of War data on war: An update to 1997. *Conflict Management and Peace Science* 18 (1): 123-4

Savun, B., Tirone, D. (2011). Foreign aid, democratization, and civil conflict: How does democracy aid affect civil conflict? *American Journal of Political Science*, 55, 233–246.

Schlumberger, O. (2021). *Puzzles of Political Change in the Middle East :Political Liberalization, Authoritarian Resilience, and the Question of Systemic Change*. Discussion Paper N° 05/2021. Deutsches Institut für Entwicklungspolitik (DIE) ; Germany : Bonn.

Steel M. F. J. (2020). “Model Averaging and Its Use in Economics”. *Journal of Economic Literature,* 58(3), 644-719.

Walter, B.F. (2017), “The New New Civil War”, *Annual Review of Political Science*, 20, 469-486.

Ward, Michael D., and Kristian S. Gleditsch. (2002). "Location, Location, Location: An MCMC Approach to Modeling the Spatial Context of War and Peace." *Political Analysis* 10(3): 244-26

World Bank (2017). Beyond Scarcity: Water Security in the Middle East and North Africa, MENA Development Series (World Bank: Washington, DC).

Yesilyurt, M Ensar, and J Paul Elhorst. (2017). “Impacts of Neighboring Countries on Military Expenditures: A Dynamic Spatial Panel Approach.” *Journal of Peace Research* 54 (6): 777–90.

On the Political Economy of Conflicts in the Middle East and Africa

Appendix

Olivier Parent

Department of Economics

Carl H. Lindner College of Business

University of Cincinnati

Cincinnati, OH, 45221

olivier.parent@uc.edu

And

Abdallah Zouache

Sciences Po Lille

9, rue Angelier, 59000 Lille, France

abdallah.zouache@sciencespo-lille.eu

Appendix A – Bayesian Model Averaging for Ordered Probit Models

We now describe the implementation of BMA using the Markov Chain Monte Carlo Model Composition (MC^3) approach. To measure the intensity of armed conflict we stratify observations into discrete categories. Ordinal conflict measures dissociate the state of peace from different levels of violent conflicts (Besley and Persson, 2009).

The latent variable $z_{i,t}$ represents the propensity of a country *i=1,…,n* at time *t=1,…,T* for entering conflict intensity *j=0,…,J-1*:

$$z_{i,t} = \alpha + x_{i,t}\gamma + \sum_{l=0,l\neq i}^{n} w_{il}x_{l,t}\theta + e_{i,t}$$

$$y_{i,t} = \mathrm{j} \quad if \;\; \delta_j < z_{i,t} \leq \delta_{j+1}, \qquad j = 0, \dots, J-1$$

where the cut-off points δ_j are unknown and for identification purposes we set $-\infty = \delta_0 < \delta_1 < \cdots < \delta_{J-1} < \delta_J = +\infty$ and $\delta_1 = 0$. The $nT\times 1$ vector Y of ordered categorical outcomes corresponds to the observed level of conflict intensity and x is the $nT\times k$ matrix of covariates. Each response $y_{i,t}$ takes the value *j=(0,…,J-1).* The probability that country *i* is involved in a conflict of intensity *j* at time *t* corresponds to $p_{i,j,t} = P(y_{i,t} = j)$. A data augmentation approach is pursued to evaluate each probability $p_{i,j,t}$ (Albert and Chib, 1993). The correspondence between $z_{i,t}$ and $y_{i,t}$ relies on different boundaries that reflect the natural ordering of the outcome. The latent variable $y_{i,t}$ will be generated from a truncated normal distribution.

The error term $e_{i,t}$ follows a standard normal distribution with mean *0* and the variance σ_e^2 is set to one to ensure the model is identifiable. Each coefficient w_{ij} of the $n\times n$ spatial weight matrix W is equal to one if countries i and j share a common border and zero otherwise. W is row-normalized so that each spatial lag $Wx_{k,t}$ represents the average of the neighboring values for the variable $x_{k,t}$. Let $X=[x\ Wx]$ be the covariates matrix of dimension $nT\times 2k$ and $\beta=(\gamma',\theta')'$ the p-dimensional vector of interest, with $p=2k$. The matrix X includes all covariates x and their neighboring effects Wx. Henceforward, we assume that X is centered.

The selection of determinants is achieved by introducing a p-dimensional vector η whose jth element is either 1 if the jth variable is included or 0 otherwise. In its simplest form, the prior distribution for η is defined as $p(\eta|\omega) = \omega^{p_\eta}(1-\omega)^{p-p_\eta}$, where p_η represents the number of selected covariates and ω is the proportion of covariates thought to be related with the outcome a priori. This proportion being unknown, it is often recommended to add a Beta hyperprior on ω instead of making an arbitrary choice.

A vague prior is assigned for the intercept $\alpha \sim N(\alpha_0, \sigma_\alpha^2)$ by setting a large variance σ_α^2 . The marginal effects β_η of the included variable follow a Normal prior distribution. As discussed in Brown et al. (2002), the conjugate prior $\beta_\eta \sim N\left(\beta_{0\eta}, H_\eta\right)$ with $H_\eta = cI_\eta$ is easier to calibrate as opposed to the traditional Zellner g-prior. The precision parameter (1/c) acts as a ridge parameter and can regulate the amount of shrinkage. The parameter c should be set such that the relative precision of the ratio prior to posterior is relatively small. As for the cutoff points δ_j, we follow Albert and Chib (1993) and assign diffuse priors using uniform distributions on each interval $(\delta_{j-1}, \delta_{j+1})$. Posterior inference is performed using Markov Chain Monte Carlo model composition (MC^3).

By integrating out the parameters α and β, the sampling procedure is simply based on the following three steps:

1. Update the latent variable $z_{i,t}$ from its posterior distribution $p(z|\eta,\delta,y)$ defined as:

$$z|\eta,\delta,y \ \sim N_{\delta}\big(\iota_{nT}\alpha_0 + X_\eta \beta_{0\eta}, \Omega_\eta\big),$$

where Ω_η=$I_{nT} + \sigma_\alpha^2 \iota_{nT}\iota_{nT}{}' + \mathrm{c}X_\eta X_\eta{}'$, where $\iota_{nT}$is an nT-dimensional vector of ones and I_{nT} is the $nT \times nT$ identity matrix.

2. Update the selection vector η using a random walk chain Metropolis-Hastings step. The conditional posterior distribution is defined as:

$$p(\eta|\delta,y,z) \propto p(\eta)p(z|\eta,\delta,y)$$

$$\propto p(\eta)\left|I_{nT} + \sigma_\alpha^2 \iota_{nT}\iota_{nT}' + cX_\eta X_\eta'\right|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left[-(nT\bar{z})^2\left(\frac{1}{nT+\frac{1}{\sigma_\alpha^2}}\right)+\left(z-\widetilde{X_\eta}\hat{\beta}_\eta\right)'\left(z-\widetilde{X_\eta}\hat{\beta}_\eta\right)\right]\right\}$$

where $\hat{\beta} = \left(\widetilde{X_\eta}\widetilde{X_\eta}'\right)^{-1}\widetilde{X_\eta}'z$, with $\widetilde{X_\eta} = X_\eta c^{1/2}$, and $\bar{z}$ represents the mean of z (see Brown et al., 2002 for further details). A candidate vector η^* is generated from the current η by one of the following transition moves. Either adding or removing a variable by changing one element of η or swapping an included with an excluded covariate selected at random.

3. Update the cut-off point parameters δ_j from their conditional posterior distribution:

$$\delta_j|\delta_{-j},\eta,z,y \sim U[a_j,b_j]$$

where a_j=max$\{\delta_{j-1}, \max_{i,t:y_{i,t}=j} z_{i,t}\}$ and b_j=min$\{\delta_{j+1}, \min_{i,t:y_{i,t}=j+1} z_{i,t}\}$.The MC^3algorithm results in a list of unique models with their corresponding relative posterior probabilities. Unlike the traditional linear model, the ordered probit requires the estimation of the latent variable z and the boundary parameters δ as described in step 1 and step 3 of the proposed MC^3 algorithm. For each model visited, the normalized conditional probabilities p$(\eta|\hat{z},\hat{\delta},y)$ are computed by averaging over the sampled z and

δ in order to obtain $\hat{z}$ and $\hat{\delta}$, respectively. With a similar approach, we derive the marginal probability of inclusion for each covariate $p(\eta_k = 1|\hat{z}, \hat{\delta}, y)$.

Appendix B – Predictive inference with Bayesian Model Averaging

Predictive inference for ordered probit models evaluates the probability that a country *i* at time *t* would be involved in any category of conflicts intensity. Instead of relying on a single model, Bayesian model averaging evaluates $\hat{z}$ over a set of selected models weighted over their posterior model probability:

$$\hat{z} = \sum_{\eta} \left(\iota_{nT}\hat{\alpha} + X_{\eta}\tilde{\beta}_{\eta}\right) p\left(\eta \middle| \hat{z}, \hat{\delta}, y\right),$$

where $\hat{\alpha} = \bar{\bar{z}}$, $\tilde{\beta}_{\eta} = \left(X'_{\eta}X_{\eta} + H_{\eta}^{-1}\right)^{-1} X'_{\eta}\hat{z}$ and $H_{\eta}=cI_{\eta}$. Different categories of conflict intensity can then be predicted for each country using:

$$\hat{y}_{i,t} = \mathrm{j} \quad if \quad \hat{\delta}_j < \hat{z}_{i,t} \leq \hat{\delta}_{j+1}, \qquad j = 0, \dots, J-1$$

To assess the predictive accuracy of the proposed method, we implement a leave-one-out cross-validation (LOO-CV) method. The conditional predictive distribution $p(y_{i,t} = \mathrm{j} \mid y_{(-i),t})$ for a country *i* at *t* to belong to the category *j* of conflict intensity is obtained by removing the *i*-th observation. LOO-CV is approximated using importance sampling (Gelfand, Dey and Chang 1992). Using a subset *s=(1,…,S)* of the MCMC draws, the conditional predictive distribution is approximated implementing $p(\eta, \delta, z|y)$ as the importance function:

$$p(y_{i,t} = \mathrm{j} \,|y_{(-i),t}) = \int_\eta \int_\delta \int_z p(y_{i,t} = \mathrm{j} \,|y_{(-i),t}, \eta, \delta, z) p(\eta, \delta, z | y_{(-i),t}) \; dz \, d\delta \, d\eta$$
$$\propto \frac{1}{S} \sum_{s=1}^{S} p\left(\delta_j^{(s)} < z_{i,t} \leq \delta_{j+1}^{(s)} \middle| y_{(-i),t}, \eta^{(t)}, z^{(s)}\right)$$
$$= \frac{1}{S} \sum_{s=1}^{S} \Phi\left(\delta_{j+1}^{(s)} - \alpha^{(s)} - x_{i,t,\eta^{(s)}} \beta_{\eta^{(s)}}\right) - \Phi\left(\delta_j^{(s)} - \alpha^{(s)} - x_{i,t,\eta^{(s)}} \beta_{\eta^{(s)}}\right)$$

where for each country *i* at time t, $x_{i,t,\eta^{(s)}}$ corresponds to the factors selected via the vector $\eta^{(s)}$ and $\Phi(.)$ represents the cumulative normal density function where $\alpha^{(s)} = \bar{z}$ and $\beta_{\eta^{(s)}} = \left(X'_{\eta^{(s)}} X_{\eta^{(s)}} + H^{-1}_{\eta^{(s)}}\right)^{-1} X'_{\eta^{(s)}} z$ are obtained by removing the *i*-th observation from the full posterior.

To predict the category *j* of conflict intensity a country belongs to, we use the mode of the predictive distribution:

$$\hat{y}_{i,t} = \underset{0 \leq j \leq J-1}{\operatorname{argmax}} \, p(y_{i,t} = \mathrm{j} \,|y_{(-i),t}).$$

We compare the proposed MC3 algorithm with more common classification methods, namely Linear Discriminant Analysis (LDA), k-Neared Neighbor (KNN) and Support Vector Machine (SVM), which do not control for the natural ordering of the different conflicts intensities. Those selection algorithms have been discussed at length in the data science literature (Duda et al., 2002). To increase predictive power, we train 100 weak classifiers for the LDA and KNN algorithms and 10 binary learners with Gaussian kernel for the SVM method. Each model is then trained using *nT -1* observations reserving one observation for validation. As those classification methods do not perform variable selection, we use the entire set of predictors. For each type of classifiers, we only report the best performing combination of parameters. We use $k=7$ for the KNN approach as it provides the largest prediction accuracy.

Results of comparative tests are presented in Table 4. With misclassification rates around 55%, classifiers perform better than the 80 percent chance of misclassification for random prediction. Our proposed method exceeds by 20% that of classification method. Finally, prediction accuracy sharply deteriorates when relying on a single order probit model even if it corresponds to the best specification.

Appendix C – Bayesian Imputation for Missing Data

A variety of Bayesian model selection procedures have been trying to unify the selection mechanism with the handling of missing data (Yang, Belin, and Boscardin, 2005). Those approaches imbedding the imputation step have mainly been developed for the stochastic search variables selection method.

We separate both processes and implement a data augmentation step to impute missing values (Gelman et al., 2004). First, we start our proposed MC3 procedure ignoring the missing variable problem. Then, for each covariate containing missing values, we select the best model this covariate belongs to and implement the following data augmentation approach where imputed data are filled in for the missing values. Table 3 compares summary statistics between the observed dataset containing missing information and the dataset replacing missing values with imputed data. The entire BMA procedure is then run again using the new imputed dataset. It is important to note that the original dataset is only missing less than 10% of its observations. To simplify the imputation process, the observed dependent variable y is assumed continuous such that the regression model can be rewritten as:

$$y|x, \beta_\eta, \sigma_e^2 \sim N(X\beta_\eta, \sigma_e^2 I_{nT}).$$

The intercept is included in the matrix of covariates. Let X_k^{mis} and X_k^{obs} denote the vectors of missing and observed values for each partially observed nT-dimensional covariate X_k. Many covariates do not have missing elements. Fully observed covariates are denoted by the $(nT \times q)$-dimensional matrix Z which is a subset $(nT \times p)$-dimensional matrix X, with $q<p$. For each covariate X_k that needs imputation, we use the set Z of observable covariates, and we make the following distributional assumption

$X_k \sim N(Z\theta, \sigma_\theta^2 I_{nT})$.

We assume priors of the form $\theta \sim N(v_0, V_0^{-1})$, $\beta_\eta \sim N(\beta_{0\eta}, H_\eta)$, $\sigma_\theta^2 \sim IG(a_1, b_1)$, and $\sigma_e^2 \sim IG(a_2, b_2)$. Each missing component $X_{i,k}^{mis}$ is then generated from the following conditional posterior distribution:

$$X_{i,k}^{mis} |\, y_i, \beta_\eta, \sigma_e^2, \theta, \sigma_\theta^2 \sim N\left(Z_i\theta + \frac{\beta_k \sigma_\theta^2}{\sigma_e^2 + \beta_k^2 \sigma_\theta^2}[y_i - \beta_k(Z_i\theta)], \frac{\sigma_e^2 \sigma_\theta^2}{(\sigma_e^2 + \beta_k^2 \sigma_\theta^2)}\right)$$

The remaining parameters related to the imputation process are obtained from the following posterior distributions:

$$\theta |\, \mathrm{y}, \beta_\eta, \sigma_\theta^2 \sim N\left(\left(Z'Z\sigma_\theta^{-2} + V_0^{-1}\right)^{-1}(Z'y\ \sigma_\theta^{-2} + V_0^{-1}v_0), \left(Z'Z\sigma_\theta^{-2} + V_0^{-1}\right)^{-1}\right)$$

$$\sigma_\theta^2 |\, X_k^{mis}, X_k^{obs}, \mathrm{y}, \theta \sim IG\left(\frac{nT}{2} + a_1, \left[b_1^{-1} + \left(\frac{1}{2}\right)(X_k - Z\theta)'(X_k - Z\theta)\right]^{-1}\right),$$

The imputed covariates are now used to draw inference on the remaining parameters:

$$\beta_\eta |\, X_k^{mis}, X_k^{obs}, \mathrm{y}, \sigma_e^2 \sim N\left(\left(X'X\sigma_e^{-2} + H_\eta^{-1}\right)^{-1}(X'y\sigma_e^{-2} + H_\eta^{-1}\beta_{0\eta}), \left(X'X\sigma_e^{-2} + H_\eta^{-1}\right)^{-1}\right),$$

$$\sigma_e^2 |\, X_k^{mis}, X_k^{obs}, \mathrm{y}, \beta_\eta, \sigma_\theta^2 \sim IG\left(\frac{nT}{2} + a_2, \left[b_2^{-1} + \left(\frac{1}{2}\right)(y - X\beta_\eta)'(y - X\beta_\eta)\right]^{-1}\right).$$

FIGURES

Figure 1. Conflict fatalities over the period 1989-2018 (in log)

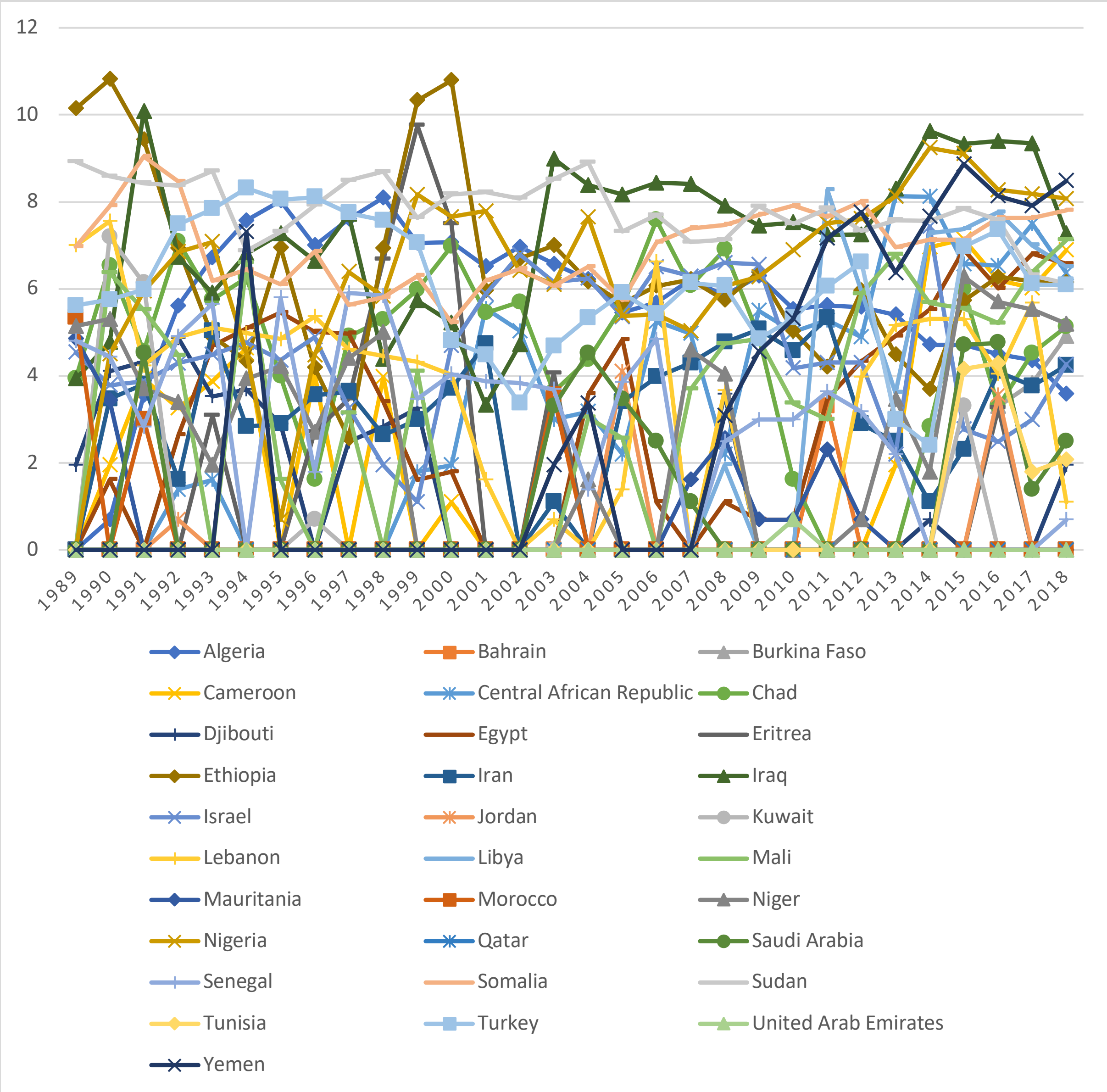


Figure 2. Location of conflict-related fatalities over the period 1989-2018

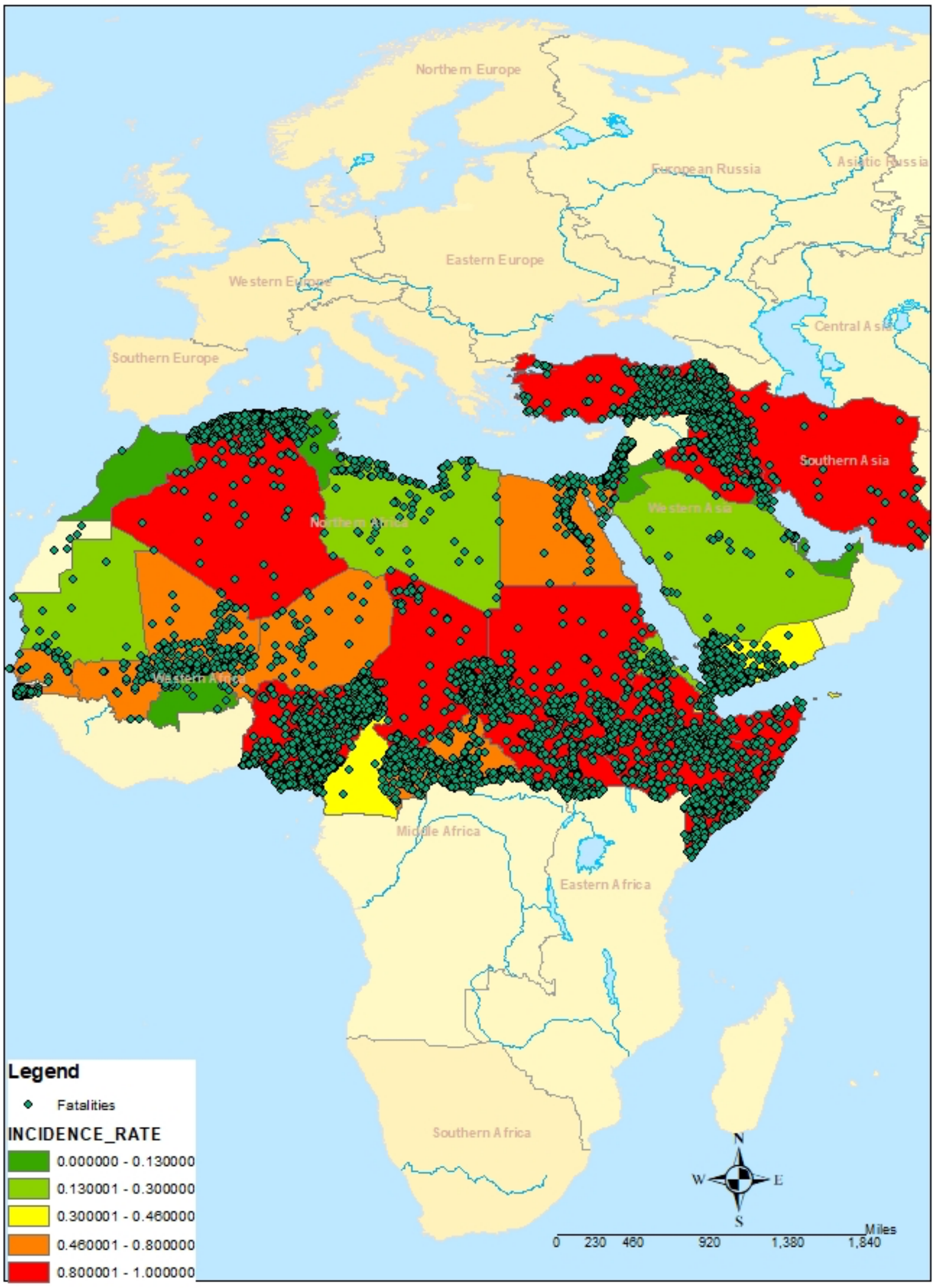


Figure 3. Average conflict intensities

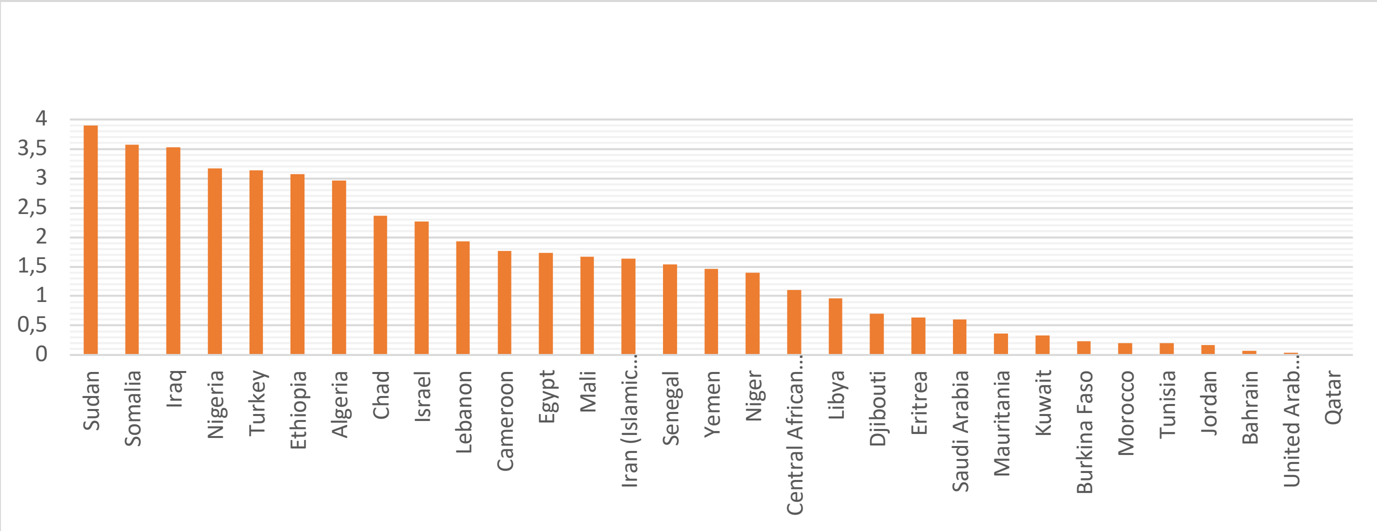
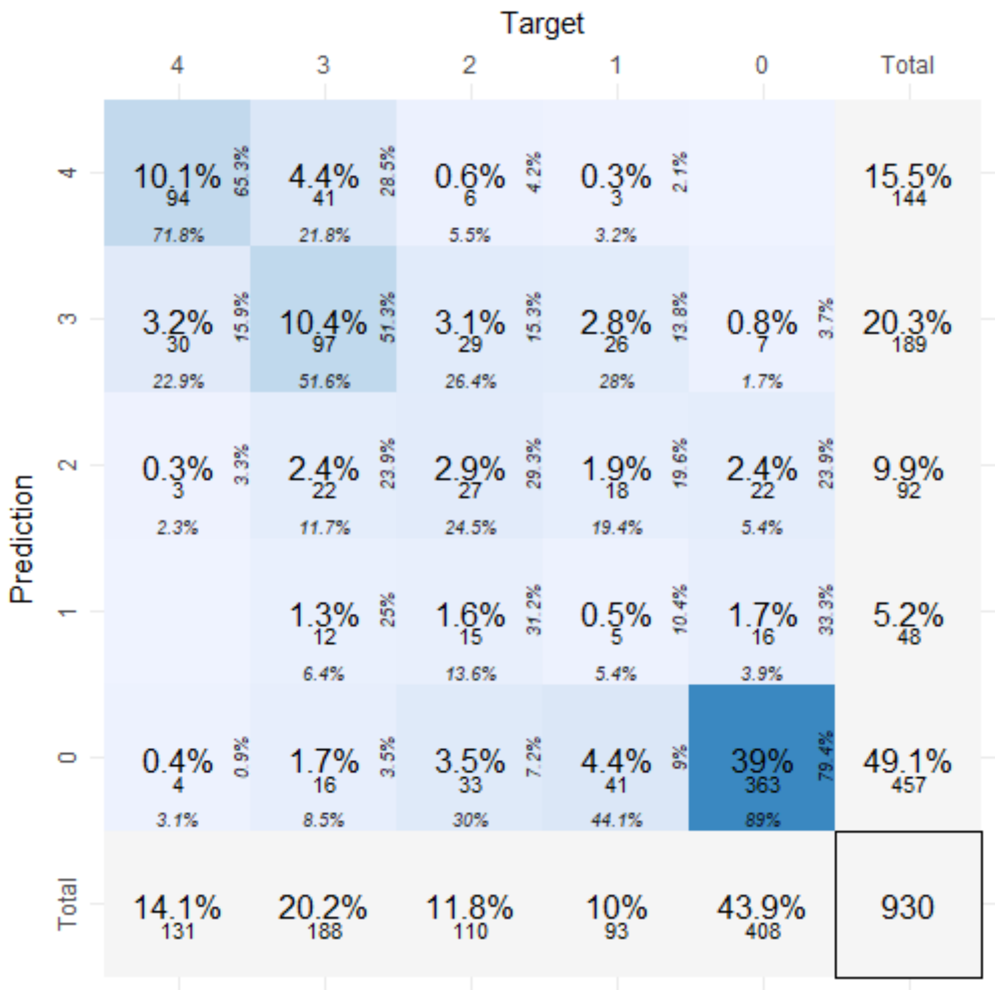


Figure 4. Confusion matrix for leave-one-out cross validation


Target
Prediction
4
3
2
1
0
Total
10.1% 94 65.3% 71.8%
4.4% 41 28.5% 21.8%
0.6% 6 4.2% 5.5%
0.3% 3 2.1% 3.2%
15.5% 144
3.2% 30 15.9% 22.9%
10.4% 97 51.3% 51.6%
3.1% 29 15.3% 26.4%
2.8% 26 13.8% 28%
0.8% 7 3.7% 1.7%
20.3% 189
0.3% 3 3.3% 2.3%
2.4% 22 23.9% 11.7%
2.9% 27 29.3% 24.5%
1.9% 18 19.6% 19.4%
2.4% 22 23.9% 5.4%
9.9% 92
1.3% 12 25% 6.4%
1.6% 15 31.2% 13.6%
0.5% 5 10.4% 5.4%
1.7% 16 33.3% 3.9%
5.2% 48
0.4% 4 0.9% 3.1%
1.7% 16 3.5% 8.5%
3.5% 33 7.2% 30%
4.4% 41 9% 44.1%
39% 363 79.4% 89%
49.1% 457
14.1% 131
20.2% 188
11.8% 110
10% 93
43.9% 408
930

TABLES (appendix or not?)

Table 1. Variable description and summary statistics

The first category of conflict determinants contains the political economy variables presented in the theoretical literature reviewed in section 3. The second category encompasses all the demographic and geographical factors including climatic variables that have been considered in the vast empirical literature on conflict (Fearon and Laitin 2003; Sambanis 2004; Lacina, 2006; Gleditsch 2021). Summary statistics for all potential determinants of conflicts are presented in Table 1.

The political economy literature highlights the role of institutional, historical, cultural, socio-economic, geographic, and environmental variables.

Institutions

As previously mentioned, political and economic institutions play a central role in preventing conflict onset. Besley and Persson (2009) underline that weak state capacity can lead to higher risk of conflict. The lack of ability for a state to collect taxes, provide public goods and enforce law provides a conducive environment for rent-seeking behaviour, corruption and warfare. Using multiple existing data sets (Freedom House, Polity Project …) assembled by the Vdem institute, we consider a series of institutional variables measuring government accountability, civil liberties, and freedom of press, academic and cultural expression. A neopatrimonial index is also considered since the literature insists on the impact of political organization on economic stability in the Arab countries (Diwan et al., 2019). Finally, we rely on another recent scholarly debate over the impact of foreign aid on the stability of the democratization process (Savun and Tirone, 2011). We implement a set of factors related to development assistance (OECD, 2021) and countries' voting affinity in the United Nations General Assembly (Bailey et al., 2017.

Whereas many studies have found no significant relationship between militarization and risk of conflict (Suzuki, 2007), military spending of neighboring countries is often related to concerns over regional instability. In terms of military expenditure, a difference should be made between Africa and the Middle East. The SIPRI (2022) estimates that, in 2021, Africa accounts for 1.9% and the Middle East for 8.8% of military expenditure worldwide. Differences also exist between countries with Saudi Arabia, Israel, and Iran being all listed among the top 20 spenders, whereas Algeria is the leading African country (SIPRI, 2022). We also include a set of variables based on armed force personnel and armed imports from the World Bank dataset.

Historical variables

The negative relationship between early statehood and contemporary economic development has been extensively analysed in economics (Bocksette et al. 2002) and political science (Hariri, 2012), emphasizing that older autocratic regimes did not benefit from European institutions. Olsson and Paik (2016) argue that old civilizations developed highly centralized, extractive and autocratic institutions, which impeded the emergence of democratization and technological innovation.

Based on the Borcan et al. (2018) dataset, we calculate a series of state history scores separating the precolonial period 3500 BCE – 1450 CE from the colonial legacy starting in 1450 CE with the Portuguese explorers and settlers.

Decisions about borders have been made by colonial powers on lines drawn up by rent-seeking Leviathans often ignoring the preference of the different local ethnic, cultural, linguistic and religious groups. The impact of artificial border and ethnic fractionalization has been discussed at length in Alesina et al. (2011). Mostly based on their database, we introduce a set of variables capturing territorial and ethnic diversity.

In Northern Africa and the Middle East, countries were mainly under the control of two principal former colonial powers, Britain and France. Italy was a smaller colonial power extending its control over Libya and the Horn of Africa, two regions that suffered extensively from civil wars. We test this colonial legacy using dichotomous variables.

Cultural variables

A large set of variables are introduced to capture cultural and religious discourses and practices that are often seen as at the heart of the political economy problems in the Arab world, and as impeding economic development (Kuran, 2011). Norms, laws, customs and practices which govern the decisions of social and economic groups and, ultimately, individual decisions would depend on the structure of beliefs carried by a culture which is conceived above all as a religious filter (Greif 1994). Culture explains the formation of different beliefs that build particular institutions (conflict resolution by law, securing property) that are the breeding ground for the emergence of organizational structures (corporations, businesses) whose economic development depends on advances in technical knowledge and accumulation of human capital. Lewis (1964) argues that Islamic culture, in both its Arabic and Turkish variants, is incompatible with democracy and favours 'military virtues'. Thus, authoritarian states arise rather than democratic nation-states that would suppose the abandonment of tribal cultural values. Concerning the region we study, many scholars have insisted on the opposition between Shias and Sunnis as key to the understanding of current and future conflicts (Nasr, 2006). We have thus included a set of variables reflecting the degree of religious pluralism.

Demographic and socio-economic factors

The extensive literature addressing the relationship between demographic factors and conflict has emphasized that countries in earlier stages of demographic transition have a greater risk of conflict. Those countries are typically characterized by large populations of young adults, rapid urban population growth rates, and high rates of infant mortality. We use the following demographic factors from the World Bank dataset: population density, percentage of population aged 14 and under, infant and adult mortality rates, growth rate of rural population, and percentage of rural population.

Similarly, countries with low economic development are more likely to be involved in a conflict (Collier and Hoeffler 2004). For this reason, we consider both the level and growth rate of GDP per capita. We also include variables related to international trade to examine whether trade and foreign investment reduce conflict propensity. Among other standard economic variables, we control for debt and employment. Given the structural characteristics of these economies, and particularly the role of natural resources and the relatively high share of agriculture in the countries' composition of GDP and employment, we also incorporate a series of variables related to agriculture and the rentier system.

Educational factors are included in the form of the share of population in primary and secondary education. The data are obtained from the World Bank education data set (WDI education). We also add a variable that measures the share of uneducated persons under the age of 14, to control for the impact of the lack of educational attainment in countries that are known for the large percentage of young in their population.

Geographical and environmental variables

Because all countries are located in northern Africa and the Middle East, we constrain the geographical factors to mainly climate factors in order to capture differences in arable land and agricultural intensity. Countries located in arid regions such as Bahrain and Kuwait have a surface area of arable land per person more than 100 times smaller than tropical countries like Niger, Central African Republic or Burkina Faso. Similar observations can be drawn for the percentage of agricultural land, which is more than 10 times smaller in Egypt, UAE or Libya than it is in Nigeria, Morocco or Tunisia. Countries with scarce resources, which experience rapid population growth, face inequitable access to arable land and renewable fresh water. Drawn from the Food Agriculture Organization Aquastat database (2019), a series of variables related to the access to safe drinking water, agricultural area equipped with irrigation, water inflow, dam capacity and level of desalination measure the impact of freshwater sustainability on political stability. Focusing on the association between water accessibility and human use, we implement different water stress indexes from the World Resource Institute (Hofste et al., 2019). The historical average of drought length from 1901 to 2008 and the seasonal variability of water supply capture the impact of extreme weather events.

Category	Variable	Mean	Std.Dev.	Description	Source
Institutional	Affinity_China	0.725	0.321	Affinity score towards China based on Votes in the United Nations General Assembly (UNGA) – Version 27 (April 29, 2020)	Voeten et al. (2009)
	Affinity_Russia	0.567	0.265	Affinity score towards Russia based on Votes in the United Nations General Assembly (UNGA) – Version 27 (April 29, 2020)	Voeten et al. (2009)

Affinity_USA	0.144	0.134	Affinity score towards USA based on Votes in the United Nations General Assembly (UNGA) – Version 27 (April 29, 2020)	Voeten et al. (2009)
Govt_Effectivness	-0.518	0.759	Index measuring quality of public services, policy implementation, and the credibility of the government's commitment (from -2.5= weak, to 2.5=strong performance).	Worldwide Governance Indicators
Academic_Free	2.004	0.922	Freedom of academic and cultural expression, from not respected to fully respected by public authorities (0-4)	University of Gothenburg, V-Dem Institute
Accountability	0.497	0.25	Government accountability index, from low to high (0-1)	University of Gothenburg, V-Dem Institute

Civil_Lib	0.466	0.215	Civil liberties index, from low to high (0-1)	University of Gothenburg, V-Dem Institute
Client	1.611	0.74	Party linkages to constituents, from clientelistic to policy driven (0-4)	University of Gothenburg, V-Dem Institute
Corrpt	1.541	0.791	Executive embezzlement and theft, from constantly to never (0-4)	University of Gothenburg, V-Dem Institute
Free_Assoc	0.426	0.282	Freedom of association thick index, from low to high (0-1)	University of Gothenburg, V-Dem Institute
Health_Equ	1.852	0.889	Educational equality, from extreme to equal (0 to 4)	University of Gothenburg, V-Dem Institute
Media_Free	1.591	0.871	Media censorship effort, from routine to rarely (0-4)	University of Gothenburg, V-Dem Institute

Neopatron	0.683	0.212	Neopatrimonial Rule Index based on Clientelism, Presidentialism and Regime Corruption, from low to high (0-1)	University of Gothenburg, V-Dem Institute
ODA_Commit	0.59	1.515	Official Development Assistance (ODA) - Total commitments	Organisation for Economic Cooperation and Development
Polyarchy	0.3	0.194	Electoral democracy index, from low to high (0-1)	University of Gothenburg, V-Dem Institute
religi_free	2.307	0.937	Freedom of religion, from not respected to fully respected by public authorities (0-4)	University of Gothenburg, V-Dem Institute
Rule_Law	0.356	0.232	Rule of law index, from low to high (0-1)	University of Gothenburg, V-Dem Institute

	Fragile_State_Change	0.337	0.971	Yearly change for Fragile State Index	Fund for Peace
	Fragile_State_Mean	85.633	16.607	Fragile State Index, 0 (low) - 120 (high risk)	Fund for Peace
	Past_3_Years_Conflict	0.459	0.499	Whether the country was involved in a conflict over the past three year (1=yes; 0=no)	Authors
	Perc_Confl_Independ	0.251	0.266	Years of conflicts since independence (%)	Authors
Socio-Economic	Access_Electr	45.916	41.495	Access to electricity (% of population)	World Bank
	Acct_Bal	1.691	15.858	Current account balance (BoP, In Billion Current US Dollars)	World Bank
	Adj_Savings_Educ	3.234	1.85	Adjusted savings: education expenditure (% of GNI)	World Bank
	Agri_Val_Add	14.122	14.678	Agriculture, forestry, and fishing, value added (% of GDP)	World Bank

Agri_Land	32.408	25.233	Agricultural land (% of land area)	World Bank
Arm_Force_Perc	2.545	3.06	Armed forces personnel (% of total labor force)	World Bank
Arm_Personnel	0.144	0.207	Armed forces personnel, in Million	World Bank
Arm_Imp	0.236	0.494	Arms imports (SIPRI trend indicator values, In Billion Current US Dollars)	World Bank
Educ_Equal	1.705	0.787	Health equality, from extreme to equal (0 to 4)	University of Gothenburg, V-Dem Institute
Empl_Pop_Fem	33.202	22.252	Employment to population ratio, 15+, female (%) (modeled ILO estimate)	World Bank
Empl_To_Pop	51.2	20.436	Employment to population ratio, 15+, total (%) (modeled ILO estimate)	World Bank
FDI	1.449	3.691	Foreign direct investment, net	World Bank

			inflows (BoP, In Billion Current US Dollars)	
FDI_GDP	2.348	4.251	Foreign direct investment, net inflows (% of GDP)	World Bank
GDPgr	3.988	8.101	GDP growth (annual %)	World Bank
GDPpcgr	1.352	7.789	GDP per capita growth (annual %)	World Bank
Gender_Ineq	0.418	0.23	Gender Inequality Index (GII) [equality = 0; inequality = 1)	United Nations Development Programme
Merchan_Exp_MENA	6.201	10.015	Merchandise exports to low- and middle-income economies in Middle East & North Africa (% of total merchandise exports)	World Bank
Metal_Exp	6.034	13.85	Ores and metals exports (% of merchandise exports)	Natural Resource Governance Institute

Mineral_rent	0.96	4.12	Mineral rents (% of GDP)	Natural Resource Governance Institute
Mort	229.566	123.66	Mortality rate, adult, male (per 1,000 male adults)	World Bank
Mort_Inf	48.633	35.651	Mortality rate, infant (per 1,000 live births)	World Bank
Nat_Inc	59.956	116.66	Adjusted net national income (In Billion Current US Dollars)	World Bank
Nat_Res_Rent	13.482	14.78	Total natural resources rents (% of GDP)	Natural Resource Governance Institute
Pop_Density	0.106	0.238	Population density (people per sq. km of land area)"	World Bank
Pop_0-14	37.247	9.72	Population, ages 0-14 (% of total)	World Bank
Trade_Openness	59.884	39.963	Exports plus imports of goods and services (% of GDP)	World Bank

Undernourishment	6.705	11.27	Prevalence of undernourishment (%)	World Bank
Unemp	7.757	5.711	Unemployment, total (% of total labor force)	World Bank
PPG_debt	33.037	26.207	Currency composition of PPG debt, U.S. dollars (%)	World Bank
Secon_Educ	60.64	43.728	Share of all students in secondary education enrolled in general programmes (%)	World Bank
Rural_gr	1.203	8.036	Rural population growth (annual %)	World Bank
Rural_Pop	43.768	25.892	Rural population (% of total population)	World Bank
School_Age_Prim	0.529	0.742	School age population, last grade of primary education, both sexes (in Million)	World Bank
Oil_Rent	9.669	14.886	Oil rents (% of GDP)	World Bank

	Interest_Debt	0.429	1.41	Interest payments on external debt, long-term (INT, In Billion Current US Dollars)	World Bank
	Internet	11.948	21.425	Individuals using the Internet (% of population)	World Bank
	Forest_Rent	2.36	4.579	Forest rents (% of GDP)	World Bank
	External_Debt_Stocks	10.676	30.639	External debt stocks, long-term (DOD, In Billion Current US Dollars)	World Bank
	Drug_Seizures	101.894	223.56	Annual Drug Seizures (kg)	United Nations Office on Drugs and Crime
	Cereal_Yield	2.204	3.378	Cereal yield (Tons per hectare)	World Bank
Geographical	Arable_Land	0.202	0.232	Arable land (hectares per person)	World Bank

Basin_Water_Stress	-378.341	695.553	water stress index per sub-basin (0= low risk; 5=Extremely high: arid and low water use)	World Resources Institute
Border_Rivers	2.506	5.811	Surface water: total flow of border rivers (10^9 m3/year)	Food and Agriculture Organization of the United Nations
Dam_Cap_Pc	0.469	1.143	Dam capacity per capita (m3/inhab)	Food and Agriculture Organization of the United Nations
Desalination	0.103	0.328	Desalinated water produced (10^9 m3/year)	Food and Agriculture Organization of the United Nations
Drought_Risk	-2879.08	2984.91	Drought risk measures based on Carrão et al. (2016) (0= low risk; 5=Extremely high)	World Resources Institute

Perc_Irrigation	38.784	45.963	Percentage of agricultural water managed area equipped for irrigation (%)	Food and Agriculture Organization of the United Nations
Perc_Pop_Safe_Drink	68.962	30.669	Total population with access to safe drinking-water (%)	Food and Agriculture Organization of the United Nations
Perc_Rur_Safe_Drink	61.95	31.386	Rural population with access to safe drinking-water (%)	Food and Agriculture Organization of the United Nations
Seasonal_Variability	-596.392	762.481	average within-year variability of available water supply (0= low risk; 5=Extremely high)	World Resources Institute
Water_Entering	14.751	27.688	Surface water: entering the country (total) (10^9 m3/year)	Food and Agriculture Organization of the United Nations

Water_risk_Agri	3.489	0.438	Risk associated with total annual agricultural water withdrawals (0= low risk; 5=Extremely high)	World Resources Institute
Outflow	14.641	31.164	Surface water: outflow to other countries not submitted to treaties (10^9 m3/year)	Food and Agriculture Organization of the United Nations
Landlock	0.194	0.395	Whether the country is landlocked (1=Yes; 0=No)	Authors
Renewed_water_pc	2.389	6.726	Total internal renewable water resources per capita (m3/inhab/year)	Food and Agriculture Organization of the United Nations
Inflow_Treaties	2.962	10.77	Surface water: inflow secured through treaties (10^9 m3/year)	Food and Agriculture Organization of the United Nations

Historical	State_Hist_01n	0.332	0.218	Normalized aggregate state history score calculated for the period 3500 BCE - 2000 CE, discounted using 1% rate	Borcan, Olsson, Putterman (2018)
	State_Hist_1450_01n	0.58	0.276	aggregate state history score calculated for the period 3500 BCE - 1450 CE, discounted using 1% rate	Borcan, Olsson, Putterman (2018)
	State_Hist_1450_2000n	0.298	0.236	aggregate state history score calculated for the period 1450 CE - 2000 CE, discounted using 1% rate	Borcan, Olsson, Putterman (2018)
	British_Colonies	0.387	0.487	Dummy variable for former French colonies	Pew Research Center
	French_Colonies	0.419	0.494	Dummy variable for former French colonies	Pew Research Center
	Italian_Colonies	0.129	0.335	Dummy variable for former French colonies	Pew Research Center

Ethnic_Frac	0.534	0.278	Historical Index of Ethnic Fractionalization (probability that two individuals do not belong to the same ethnic group)	Drazanova (2019)
Ethno_Ling	0.304	0.277	Ethno-linguistic fractionalization index (Herfindhal Index)	Alesina, Easterly, Matuszeski (2011)
Artifical_Border	0.991	0.182	Fractal dimension of each political borders (12 boxed sizes)	Alesina, Easterly, Matuszeski (2011)
Partitioned	29.831	29.984	Share of population that belongs to a partitioned group	Alesina, Easterly, Matuszeski (2011)
Polarization	0.455	0.303	Measures the degree to which individuals are distributed across ethnic groups (Maximum with	Montalvo and Raynal Querol (1995)

		bipolar ethnic distribution).

Table 2. BMA – Estimation results

Variable	Imputation - Chain 1		Imputation - Chain 2		Imputation - Chain 3		Imputation - Chain 4	
	Prob(incl.)	Estimates	Prob(incl.)	Estimates	Prob(incl.)	Estimates	Prob(incl.)	Estimates
Intercept		1.15(***) (0.57)		1.20(***) (0.52)		1.13(***) (0.53)		1.05(***) (0.50)
Past_3_Years_Conflict	1.00	1.65(**) (0.51)	1.00	1.60(**) (0.47)	1.00	1.93(**) (0.60)	1.00	1.86(**) (0.53)
Water_risk_Agri	1.00	-2.71(***) (1.10)	1.00	-2.53(***) (1.00)	1.00	-2.86(***) (1.18)	1.00	-3.49(***) (1.43)
Desalination	1.00	1.25(**) (0.41)	1.00	1.70(**) (0.54)	1.00	1.68(**) (0.55)	1.00	1.72(**) (0.57)
Unemp	1.00	0.09(**) (0.03)	1.00	0.09(**) (0.03)	1.00	0.11(***) (0.05)	1.00	0.13(**) (0.05)
Water_Entering	1.00	0.05(**) (0.02)	1.00	0.05(**) (0.02)	1.00	0.05(***) (0.02)	1.00	0.06(**) (0.02)

W×Polarization	1.00	-1.45(**)	1.00	-1.87(**)	1.00	-1.73(**)	1.00	-1.67(**)
		(0.50)		(0.62)		(0.59)		(0.57)
Affinity_USA	1.00	2.32(***)	1.00	2.78(***)	1.00	3.48(***)	1.00	4.02(***)
		(1.18)		(1.31)		(1.55)		(1.83)
Inflow_Treaties	1.00	-0.08(***)	1.00	-0.09(***)	1.00	-0.09(***)	1.00	-0.11(***)
		(0.03)		(0.03)		(0.04)		(0.05)
Ethno_Ling	1.00	1.99(***)	1.00	1.57(**)	1.00	2.76(**)	1.00	1.21(***)
		(0.78)		(0.58)		(0.97)		(0.53)
Renewed_water_pc	1.00	-0.14(***)	1.00	-0.13(***)	1.00	-0.12(***)	1.00	-0.14(***)
		(0.06)		(0.06)		(0.05)		(0.06)
Academic_Free	1.00	-0.91(***)	1.00	-0.73(***)	1.00	-1.03(**)	1.00	-0.94(***)
		(0.38)		(0.32)		(0.38)		(0.45)
Corrpt	1.00	0.79(**)	1.00	0.65(**)	1.00	0.77(**)	1.00	0.71(***)
		(0.30)		(0.24)		(0.29)		(0.29)
Accountability	1.00	3.58(***)	0.99	4.24(***)	1.00	4.11(***)	1.00	3.81(***)
		(1.64)		(1.83)		(1.96)		(1.86)
W×Arable_Land	0.98	-1.67(**)	1.00	-1.28(**)	1.00	-1.25(***)	1.00	-2.64(**)
		(0.57)		(0.45)		(0.50)		(0.89)
Polyarchy	0.99	1.12	1.00	2.53(**)	0.97	2.57(**)	1.00	2.72(**)
		(0.77)		(0.59)		(0.86)		(0.81)
Perc_Pop_Safe_Drink	0.94	-0.03(***)	1.00	-0.02(**)	0.99	-0.03(***)	1.00	-0.08(**)
		(0.01)		(0.01)		(0.01)		(0.03)
W×Undernourishment	0.99	0.04(**)	1.00	0.04(**)	0.82	0.05(**)	1.00	0.06(**)
		(0.01)		(0.01)		(0.02)		(0.02)

Fragile_State_Change	0.42	-0.18(***)	0.96	-0.48(**)	0.99	-0.38(*)	1.00	-0.72(***)
		(0.09)		(0.18)		(0.23)		(0.30)
Govt_Effectivness	0.37	-0.42(*)	0.98	-0.49(*)	0.82	-1.36(**)	1.00	-1.65(**)
		(0.23)		(0.29)		(0.51)		(0.62)
Fragile_State_Mean	0.93	0.07(**)	0.75	0.05(***)	0.20	0.02(**)	0.84	0.04(***)
		(0.03)		(0.02)		(0.01)		(0.02)
Neopatron	0.81	0.87	0.17	0.02	0.78	1.39(*)	0.94	2.06(***)
		(0.83)		(0.01)		(0.79)		(1.03)
W×Polyarchy	0.44	0.25	0.79	-0.08	0.46	0.96(*)	0.94	1.66(***)
		(0.26)		(0.90)		(0.50)		(0.79)
W×Neopatron	0.51	0.38	0.77	2.91(*)	0.94	1.08	0.33	0.23
		(0.30)		(1.59)		(0.88)		(0.28)
GDPgr	0.52	-0.01(**)	0.84	-0.03(**)	0.76	-0.03(***)	0.33	-0.01(***)
		(0.00)		(0.01)		(0.01)		(0.00)
Italian_Colonies	0.49	-0.28(***)	0.05	-0.01(***)	0.98	-0.48(*)	0.87	-0.08
		(0.13)		(0.00)		(0.29)		(0.29)
Arm_Personnel	0.53	0.64(***)	0.76	1.15(***)	0.06	0.10(***)	0.94	1.08
		(0.28)		(0.51)		(0.05)		(0.67)
Arm_Imp	0.91	0.44(***)	0.27	0.03(***)	0.99	0.51(***)	0.05	0.03(***)
		(0.19)		(0.01)		(0.21)		(0.01)
Rule_Law	0.94	-1.43(*)	0.82	-3.16(**)	0.27	-0.56(***)	0.19	-0.49(***)
		(0.79)		(1.04)		(0.26)		(0.21)
British_Colonies	0.65	-0.70(***)	0.80	-0.83(***)	0.74	-0.33(*)	0.01	0.00
		(0.32)		(0.36)		(0.18)		(0.00)

Forest_Rent	0.39	0.04(**) (0.01)	0.93	0.09(**) (0.03)	0.06	0.01(***) (0.00)	0.80	0.08(**) (0.03)
W×Perc_Confl_Independ	0.27	0.16(***) (0.07)	0.96	0.77(***) (0.37)	0.82	0.79(***) (0.35)	0.05	0.04(*) (0.02)
Empl_Pop_Fem	0.94	-0.03(**) (0.01)	0.95	-0.04(**) (0.01)	0.18	-0.01(**) (0.00)	0.00	0.00 (0.00)
W×Ethno_Ling	0.07	-0.03(***) (0.03)	0.24	-0.03(*) (0.02)	0.69	-0.55(*) (0.28)	0.99	-1.42(***) (0.63)
W×Artifical_Border	0.06	-0.08(***) (0.04)	0.80	-2.31(**) (0.78)	0.70	-0.95(***) (0.38)	0.38	-0.38(***) (0.17)
Empl_To_Pop	0.06	0.0 (0.00)	0.05	0.00 (0.00)	0.82	-0.02(**) (0.01)	1.00	-0.04(***) (0.02)
Perc_Confl_Independ	0.41	0.24 (0.16)	0.33	-0.01 (0.07)	0.81	1.45(**) (0.42)	0.33	0.28(***) (0.14)
State_Hist_1450_2000n	0.34	6.14 (67.02)	0.76	1.43(**) (0.50)	0.03	0.63 (17.96)	0.70	21.19 (123.32)
Artifical_Border	0.63	1.24(*) (0.66)	0.04	0.01(*) (0.01)	0.76	-0.16 (0.61)	0.27	0.93(***) (0.41)
Free_Assoc	0.42	0.58 (0.50)	0.27	0.07 (0.19)	0.05	0.05 (0.05)	0.95	2.01 (1.43)
W×State_Hist_01n	0.26	-0.16(*) (0.09)	0.24	-0.04(*) (0.02)	0.06	-0.03 (0.04)	0.93	-0.78 (0.68)
State_Hist_1450_01n	0.56	0.21 (13.15)	0.75	-1.62(***) (0.77)	0.07	0.05 (3.43)	0.06	2.64 (22.78)

Pop_Density	0.40	0.44	0.01	0.02(***)	0.14	0.08	0.87	0.93(*)
		(0.29)		(0.01)		(0.10)		(0.49)
W×Affinity_Russia	0.39	-0.24(*)	0.02	0.00	0.05	-0.03	0.95	-1.14(*)
		(0.14)		(0.00)		(0.02)		(0.60)
W×State_Hist_1450_01n	0.37	0.40(*)	0.01	0.00	0.00	0.00	0.94	0.43
		(0.23)		(0.00)		(0.00)		(0.49)
Perc_Rur_Safe_Drink	0.36	0.02(***)	0.01	0.00	0.00	0.00	0.95	0.06(**)
		(0.01)		(0.00)		(0.00)		(0.02)
W×Gender_Ineq	0.26	-0.17(*)	0.19	-0.04(*)	0.79	-0.68(*)	0.05	-0.03
		(0.10)		(0.02)		(0.40)		(0.02)
Affinity_China	0.33	-0.01	0.89	-0.29	0.01	0.00	0.00	0.00
		(0.09)		(0.22)		(0.00)		(0.00)
W×Ethnic_Frac	0.07	0.04	0.03	0.02(*)	0.06	0.07(*)	0.95	1.18(*)
		(0.03)		(0.01)		(0.03)		(0.63)
W×Pop_Density	0.09	-0.02	0.78	-1.31(***)	0.24	-0.03	0.00	0.00
		(0.04)		(0.59)		(0.10)		(0.00)
Ethnic_Frac	0.01	0.00	0.22	-0.04	0.83	-1.36(***)	0.00	0.00
		(0.00)		(0.02)		(0.67)		(0.00)
W×Corrpt	0.06	-0.01	0.00	0.00	0.01	0.00	0.97	-0.41(*)
		(0.01)		(0.00)		(0.00)		(0.24)
W×Accountability	0.04	0.07	0.77	1.81	0.18	0.44(*)	0.02	0.02(*)
		(0.05)		(1.19)		(0.26)		(0.01)
W×Rule_Law	0.12	0.00	0.79	1.40(*)	0.10	0.11	0.01	0.00
		(0.04)		(0.81)		(0.10)		(0.00)

PPG_debt	0.00	0.00 (0.00)	0.20	0.00 (0.00)	0.76	0.02(***) (0.01)	0.05	0.00 (0.00)

Note: the remaining 79 variables are never selected (W×Water_risk_Agri, Interest_Debt, Pop_Density, W×Forest_Rent, W×ODA_Commit, ODA_Commit, W×Interest_Debt, W×Unemp, W×Outflow, W×Fragile_State_Mean, W×Mort, W×Renewed_water_pc, W×Adj_Savings_Educ, Undernourishment, W×Inflow_Treaties, W×Dam_Cap_Pc, Arm_Force_Perc, W×External_Debt_Stocks, Cereal_Yield, W×Arm_Force_Perc, W×Mineral_rent, Metal_Exp, W×Agri_Land, W×GDPgr, Non_Muslim, Merchan_Exp_MENA, W×FDI_GDP, Mineral_rent, Agri_Land, Nat_Res_Rent, FDI, Border_Rivers, Perc_Irrigation, FDI_GDP, W×Perc_Rur_Safe_Drink, Secon_Educ, Partitioned, W×Water_Entering, Oil_Rent, W×Oil_Rent, W×Secon_Educ, W×Nat_Inc, W×Pop_Density, W×Nat_Res_Rent, W×Non_Muslim, Trade_Openness, Access_Electr, Mort_Inf, Drug_Seizures, W×PPG_debt, W×Empl_Pop_Fem, W×Perc_Pop_Safe_Drink, Agri_Val_Add, W×Internet, W×Rural_Pop, W×Merchan_Exp_MENA, Rural_Pop, Internet, W×Partitioned, W×Acct_Bal, Acct_Bal, W×Access_Electr, W×Empl_To_Pop, Shias, W×Shias, Mort, W×Trade_Openness, W×Perc_Irrigation, W×Mort_Inf, W×Agri_Val_Add, External_Debt_Stocks, Nat_Inc, Seasonal_Variability, W×Drug_Seizures, Drought_Risk, W×Basin_Water_Stress, Basin_Water_Stress, W×Drought_Risk, W×Seasonal_Variability.)

Table 3. Summary statistics for observed data with missing information and imputed data

	Observed Data				% of	Imputed Data				
Variable	Mean	St.Dev.	Min	Max	Missing	Mean	St.Dev.	Min	Max	Frequency
Academic_Free	2.00	0.92	0.00	3.86	0.00	2.00	0.92	0.00	3.86	1989-2018
Access_Electr	45.92	41.50	0.00	100.00	0.28	63.73	33.45	2.04	100.00	1989-2018
Accountability	0.50	0.25	0.04	0.93	0.00	0.50	0.25	0.04	0.93	1989-2018
Acct_Bal	1.69	15.86	-74.40	164.76	0.30	4.22	17.36	-74.40	164.76	1989-2018
Adj_Savings_Educ	3.23	1.85	0.00	9.50	0.06	3.45	1.65	0.80	9.50	1989-2018
Affinity_China	0.72	0.32	0.00	1.00	0.14	0.84	0.14	0.11	1.00	1989-2018
Affinity_Russia	0.57	0.26	0.00	1.00	0.14	0.63	0.16	0.13	1.00	1989-2018
Affinity_USA	0.14	0.13	0.00	0.92	0.14	0.25	0.24	0.01	0.92	1989-2018
Agri_Val_Add	14.12	14.68	0.00	63.83	0.11	16.84	13.88	0.00	63.83	1989-2018

Agri_Land	32.41	25.23	0.00	80.92	0.11	34.98	23.88	0.00	100.45	1989-2018
Arable_Land	0.20	0.23	0.00	1.52	0.11	0.23	0.23	0.00	1.52	1989-2018
Arm_Force_Perc	2.55	3.06	0.00	34.89	0.11	2.99	3.04	0.07	34.89	1989-2018
Arm_Personnel	0.14	0.21	0.00	1.39	0.07	0.15	0.21	0.00	1.39	1989-2018
Arm_Imp	0.24	0.49	0.00	4.06	0.28	0.30	0.50	0.00	4.06	1989-2018
Artifical_Border	0.99	0.18	0.00	1.07	0.03	1.02	0.02	1.00	1.07	Constant
Basin_Water_Stress	-378.34	695.55	-2996.22	4.79	0.00	-378.34	695.55	-2996.22	4.79	Constant
Border_Rivers	2.51	5.81	0.00	22.00	0.00	2.51	5.81	0.00	22.00	Constant
British_Colonies	0.39	0.49	0.00	1.00	0.00	0.39	0.49	0.00	1.00	Constant
Cereal_Yield	2.20	3.38	0.00	28.13	0.10	2.50	3.41	0.08	28.13	1989-2018
Civil_Lib	0.47	0.22	0.04	0.89	0.00	0.47	0.22	0.04	0.89	1989-2018
Client	1.61	0.74	0.00	3.75	0.00	1.62	0.73	0.00	3.75	1989-2018
Corrpt	1.54	0.79	0.00	3.38	0.00	1.56	0.77	0.11	3.38	1989-2018
Dam_Cap_Pc	0.47	1.14	0.00	8.22	0.32	0.70	1.17	0.00	8.22	1989-2018
Desalination	0.10	0.33	0.00	2.18	0.10	0.16	0.35	0.00	2.18	1989-2018
Drought_Risk	-2879.08	2984.91	-8814.16	2.90	0.00	-2879.08	2984.91	-8814.16	2.90	Constant

Drug_Seizures	101.89	223.56	0.00	871.48	0.23	121.09	229.95	0.01	871.48	1989-2018
Educ_Equal	1.71	0.79	0.17	3.60	0.00	1.71	0.79	0.17	3.60	1989-2018
Empl_Pop_Fem	33.20	22.25	0.00	75.44	0.07	33.20	22.25	4.49	75.44	1989-2018
Empl_To_Pop	51.20	20.44	0.00	87.42	0.07	53.92	15.61	30.60	87.42	1989-2018
Ethnic_Frac	0.53	0.28	0.03	0.89	0.00	0.53	0.28	0.03	0.89	Constant
Ethno_Ling	0.30	0.28	0.00	0.79	0.00	0.30	0.28	0.00	0.79	Constant
External_Debt_Stocks	10.68	30.64	0.00	335.47	0.28	12.45	31.78	0.00	335.47	1989-2018

	Observed Data				% of	Imputed Data				
Variable	Mean	St.Dev.	Min	Max	Missing	Mean	St.Dev.	Min	Max	Frequency
FDI	1.45	3.69	-10.18	39.46	0.05	1.44	3.69	-10.18	39.46	1989-2018
FDI_GDP	2.35	4.25	-5.29	46.49	0.10	2.46	4.26	-5.29	46.49	1989-2018
Forest_Rent	2.36	4.58	0.00	36.07	0.11	2.82	4.59	0.00	36.07	1989-2018
Fragile_State_Change	0.34	0.97	-1.66	3.09	0.00	0.34	0.97	-1.66	3.09	Constant
Fragile_State_Mean	85.63	16.61	48.20	113.05	0.00	85.63	16.61	48.20	113.05	Constant

Free_Assoc	0.43	0.28	0.03	0.88	0.00	0.43	0.28	0.03	0.88	1989-2018
French_Colonies	0.42	0.49	0.00	1.00	0.00	0.42	0.49	0.00	1.00	Constant
GDPgr	3.99	8.10	-64.05	123.14	0.10	4.44	7.99	-64.05	123.14	1989-2018
GDPpcgr	1.35	7.79	-64.99	121.78	0.10	1.20	7.91	-64.99	121.78	1989-2018
Gender_Ineq	0.42	0.23	0.00	0.74	0.13	0.49	0.18	0.02	0.75	1989-2018
Govt_Effectivness	-0.52	0.76	-2.14	1.22	0.00	-0.52	0.76	-2.14	1.22	Constant
Health_Equ	1.85	0.89	0.17	3.60	0.00	1.85	0.89	0.17	3.60	1989-2018
Inflow_Treaties	2.96	10.77	0.00	55.50	0.00	2.96	10.77	0.00	55.50	Constant
Interest_Debt	0.43	1.41	0.00	14.75	0.29	0.55	1.43	0.80	14.75	1989-2018
Internet	11.95	21.43	0.00	98.00	0.26	13.62	21.23	0.00	98.00	1989-2018
Italian_Colonies	0.13	0.34	0.00	1.00	0.00	0.13	0.34	0.00	1.00	Constant
Landlock	0.19	0.40	0.00	1.00	0.00	0.19	0.40	0.00	1.00	Constant
Media_Free	1.59	0.87	0.00	3.75	0.00	1.60	0.86	0.06	3.75	1989-2018
Merchan_Exp_MENA	6.20	10.02	0.00	98.88	0.06	6.48	9.97	0.00	98.88	1989-2018
Metal_Exp	6.03	13.85	0.00	80.05	0.32	9.21	14.34	0.00	80.05	1989-2018
Mineral_rent	0.96	4.12	0.00	44.64	0.15	1.97	4.57	0.00	44.64	1989-2018

Mort	229.57	123.66	0.00	556.57	0.03	235.18	117.41	63.56	556.57	1989-2018
Mort_Inf	48.63	35.65	0.00	133.70	0.03	48.74	35.52	0.69	133.70	1989-2018
Nat_Inc	59.96	116.66	0.00	805.68	0.17	62.85	115.91	0.41	805.68	1989-2018
Nat_Res_Rent	13.48	14.78	0.00	68.78	0.09	15.03	14.74	0.00	68.78	1989-2018
Neopatron	0.68	0.21	0.09	0.97	0.00	0.68	0.21	0.09	0.97	1989-2018
Non_Muslim	19.60	25.08	0.00	85.00	0.00	19.60	25.08	0.00	85.00	Constant
ODA_Commit	0.59	1.51	0.00	23.54	0.13	0.66	1.51	0.00	23.54	1989-2018
Oil_Rent	9.67	14.89	0.00	67.53	0.38	13.59	14.42	0.00	67.53	1989-2018

	Observed Data				% of	Imputed Data				
Variable	Mean	St.Dev.	Min	Max	Missing	Mean	St.Dev.	Min	Max	Frequency
Outflow	14.64	31.16	0.00	141.00	0.06	14.64	31.16	0.00	141.00	Constant
Partitioned	29.83	29.98	0.00	91.10	0.35	45.07	24.48	0.00	91.10	Constant
Past_3_Years_Conflict	0.46	0.50	0.00	1.00	0.00	0.46	0.50	0.00	1.00	1989-2018
Perc_Confl_Independ	0.25	0.27	0.00	1.00	0.00	0.25	0.27	0.00	1.00	1989-2018
Perc_Irrigation	38.78	45.96	0.00	100.00	0.10	83.92	22.31	12.36	100.00	1989-2018

Perc_Pop_Safe_Drink	68.96	30.67	0.00	100.00	0.10	75.51	21.63	21.10	100.00	1989-2018
Perc_Rur_Safe_Drink	61.95	31.39	0.00	100.00	0.10	67.15	24.83	8.80	100.00	1989-2018
Polarization	0.45	0.30	0.00	0.98	0.19	0.57	0.22	0.06	1.00	Constant
Polyarchy	0.30	0.19	0.01	0.78	0.00	0.30	0.19	0.01	0.78	1989-2018
Pop_Density	0.11	0.24	0.00	2.02	0.04	0.11	0.24	0.00	2.02	1989-2018
Pop_0-14	37.25	9.72	0.00	51.89	0.01	37.49	9.18	13.08	51.89	1989-2018
PPG_debt	33.04	26.21	0.00	93.80	0.28	46.49	18.36	1.22	95.82	1989-2018
Religi_free	2.31	0.94	0.00	3.93	0.00	2.31	0.94	0.00	3.93	1989-2018
Renewed_water_pc	2.39	6.73	0.00	45.54	0.09	2.94	6.95	0.00	45.54	1989-2018
Rule_Law	0.36	0.23	0.03	0.89	0.00	0.36	0.23	0.03	0.89	1989-2018
Rural_gr	1.20	8.04	-235.79	12.99	0.03	0.97	8.22	-235.79	12.99	1989-2018
Rural_Pop	43.77	25.89	0.00	87.62	0.03	44.10	25.42	0.09	87.62	1989-2018
School_Age_Prim	0.53	0.74	0.00	4.90	0.03	0.54	0.74	0.00	4.90	1989-2018
Seasonal_Variability	-596.39	762.48	-2998.72	3.77	0.00	-596.39	762.48	-2998.72	3.77	Constant
Secon_Educ	60.64	43.73	0.00	100.00	0.34	91.71	6.85	63.50	100.00	1989-2018

Shias	13.45	25.33	0.10	95.00	0.00	13.45	25.33	0.10	95.00	Constant
State_Hist_01n	0.33	0.22	0.00	0.74	0.13	0.38	0.18	0.02	0.74	Constant
State_Hist_1450_01n	0.58	0.28	0.00	0.99	0.13	0.67	0.17	0.13	0.99	Constant
State_Hist_1450_2000n	0.30	0.24	0.00	0.79	0.19	0.35	0.20	0.00	0.79	Constant
Trade_Openness	59.88	39.96	0.00	210.16	0.14	69.46	33.32	0.02	210.16	1989-2018
Undernourishment	6.71	11.27	0.00	59.80	0.54	18.25	14.36	0.00	59.80	1989-2018
Unemp	7.76	5.71	0.00	31.84	0.07	9.89	7.95	0.14	32.00	1989-2018
Water_Entering	14.75	27.69	0.00	99.30	0.00	14.75	27.69	0.00	99.30	Constant
Water_risk_Agri	3.49	0.44	2.43	4.13	0.00	3.49	0.44	2.43	4.13	Constant

Table 4 - Misclassification rates with leave-one-out cross-validation.

Method	Misclassification rate
KNN	0.560
LDA	0.561
SVM	0.562
BMA	0.369

Best Model	0.734